



**Imagerie sismique des milieux visco-acoustiques et
visco-élastiques à deux dimensions par
stéréotomographie et inversion des formes d'ondes:
applications au champ pétrolier de Valhall**

Vincent Prieux

► **To cite this version:**

Vincent Prieux. Imagerie sismique des milieux visco-acoustiques et visco-élastiques à deux dimensions par stéréotomographie et inversion des formes d'ondes: applications au champ pétrolier de Valhall. Géophysique [physics.geo-ph]. Université Nice Sophia Antipolis, 2012. Français. NNT: . tel-00722408

HAL Id: tel-00722408

<https://theses.hal.science/tel-00722408>

Submitted on 1 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

pour obtenir le titre de

Docteur en Sciences
de l'Université de Nice-Sophia Antipolis

Spécialité : Sciences de l'Univers - Géophysique

présentée et soutenue par

Vincent PRIEUX

Imagerie sismique des milieux visco-acoustiques et visco-élastiques à deux dimensions par stéréotomographie et inversion des formes d'ondes : applications au champ pétrolier de Valhall

Thèse dirigée par **Stéphane OPERTO** et **Jean VIRIEUX**
préparée au laboratoire Géoazur, Sophia-Antipolis
financée dans le cadre du Consortium SEISCOPE

soutenue à Géoazur (Sophia Antipolis) le 18 juin 2012 à 14h devant le Jury

Olav BARKVED	Géophysicien, BP Norge AS	Invité
Hervé CHAURIS	Professeur Mines Paristech	Rapporteur
Guust NOLET	Professeur Université Nice-Sophia Antipolis	Examineur
Stéphane OPERTO	Chargé de Recherche CNRS	Directeur de Thèse
René-Edouard PLESSIX	Chercheur Senior SHELL	Examineur
Satish SINGH	Professeur Institut de Physique du Globe de Paris	Rapporteur
Nikolai SHAPIRO	Directeur de Recherche CNRS, IPGP	Examineur
Jean VIRIEUX	Professeur Université Joseph Fourier	Directeur de Thèse

Remerciements

Je dédie tout d’abord mes remerciements à mes directeurs de thèse Stéphane Operto et Jean Virieux, qui ont su me guider tout au long de ce long périple, pour avoir créé un environnement de travail dynamique et motivant, et pour m’avoir fait confiance pour mener à bien ce projet. Je remercie Stéphane Operto pour l’ensemble de ses qualités de maître de thèse, et pour avoir formé un jeune impétrant au difficile travail de recherche. J’aimerais lui reconnaître son investissement sans pareil, sa rigueur scientifique et son regard critique. Sa vision des phénomènes physiques, des problèmes numériques et son esprit de synthèse m’ont de plus éclairé un grand nombre de fois. J’exprime à Jean Virieux ma plus grande gratitude pour sa disponibilité à mes questions, pour ses corrections et pour l’ensemble de ses commentaires qui donnent une réelle *valeur ajoutée* à ce travail. J’aimerais maintenant rendre un vibrant hommage à Romain Brossier pour l’ensemble de sa contribution. Merci de nous avoir légué ton précieux travail, de ton esprit d’équipe, de ta patience lors de ces nombreuses heures au téléphone pour résoudre mes problèmes et m’insuffler ta vision, et finalement pour ton amitié. Je remercie aussi vivement Gilles Lambaré de m’avoir accueilli au sein du centre de CGG Veritas pour me former à l’outil de stéréotomographie, pour son enthousiasme, sa disponibilité et son expérience qui ont joué un rôle clé dans notre collaboration. Je me souviens aussi d’une longue discussion téléphonique un soir alors que j’étais en Master 1, grâce à laquelle je me suis lancé par la suite dans le domaine de la géophysique. Je tiens à remercier Alessandra Ribodetti pour son invariable gentillesse, ses conseils avisés sur l’atténuation, l’anisotropie, ... et pour ses attentions qui rendent la vie plus facile. Je remercie aussi les sponsors pétroliers sans qui cette thèse n’aurait pas été possible : BP, CGG Veritas, ENI, Exxon Mobil, BR Petrobras, Shell, Statoil, Saudi Aramco et Total.

Je remercie Hervé Chauris et Satish Singh d’avoir accepté d’être les rapporteurs de ce travail et du sérieux avec lequel ils ont examinés mon travail. Je remercie également O. Barkved, Guust Nolet, René-Edouard Plessix et Nikolai Shapiro d’avoir fait partie des membres de mon jury, et pour la pertinence de leurs questions lors de la soutenance.

Je remercie chaleureusement mes co-bureaux Yaser Gholami et Damien Pageot pour leur esprit d’entraide et de convivialité. Merci notamment à Yaser d’avoir été un acteur du volet anisotrope et géologique de cette thèse, et pour les duels parfois enflammés au tableau à coups de stylo Velleda. Merci également à Vincent Etienne pour ses coups de main, nos grandes discussions et notre vadrouille aux US. Je tiens aussi à remercier l’ensemble des membres qui font ou ont fait parti du groupe SEISCOPE de près ou de loin (jusqu’à Grenoble), et avec qui j’ai partagé les congrès, les *Annual Meeting* et les non moins fameuses *réunions SEISCOPE*.

J’exprime aussi ma plus vive reconnaissance aux membres de Géoazur qui ont rendu mon séjour au sein du laboratoire très agréable. Je pense aux nombreux étudiants et chercheurs avec qui j’ai partagé les repas, les pauses café et des discussions scientifiques. Je remercie

plus particulièrement le personnel informatique et administratif pour leur efficacité et leur gentillesse.

Je me dois aussi de remercier mes colocataires (ou presque) de longue et de nouvelle date : Blandine, Manu, Nico, Thibault et Sophie, ainsi que mes amis de la région et d'ailleurs qui ont rendu ma vie de thésard inoubliable. Merci également à toute ma famille pour son soutien inconditionnel : mes parents Françoise et Patrick, ainsi que mes soeurs Clémence et Hélène (avec Arnaud et Rémi). Sans oublier les deux petits nouveaux Antoine et Thomas ! Je remercie finalement Flore avec qui j'ai partagé tellement de choses et qui m'a tant apporté.

Enfin, je remercie toutes les personnes que j'aurai pu oublier de ne pas m'en vouloir.

Résumé

La géophysique interne est une discipline riche en enseignements sur la structure de la Terre à différentes échelles, et sur les phénomènes passés liés à sa lente évolution. Elle constitue de plus un enjeu présent et d'avenir de première importance dans notre société, à l'heure où les ressources de toutes sortes se font de plus en plus rares, et dans le cadre d'une meilleure gestion de l'aménagement du territoire. La méthode d'inversion des formes d'ondes complètes, fondée sur un processus d'optimisation local, se distingue des autres méthodes d'imagerie sismique par sa vocation à extraire de manière aussi complète que possible l'intégralité de l'information sismique, afin de construire une image quantitative haute résolution d'un ou plusieurs paramètres physiques. Je discute dans ma thèse de plusieurs problématiques liées à cette méthode, en les illustrant par des applications à des données multicomposantes enregistrées par des câbles de fond de mer (OBC) sur le champ pétrolier de Valhall en mer du nord.

Je discute tout d'abord de l'empreinte de l'anisotropie engendrée par des milieux transverses isotropes à axe de symétrie vertical sur les données du champ de Valhall. J'illustre cette empreinte sur les résultats de l'imagerie en comparant les modèles du sous-sol obtenus lorsque l'anisotropie est prise en compte ou pas dans la modélisation sismique. Je mets en évidence un biais dans les vitesses reconstruites par une méthode d'inversion isotrope, ce biais induisant un mauvais positionnement des réflecteurs en profondeur.

J'aborde ensuite le problème de la construction du modèle initial nécessaire à l'inversion des formes d'ondes à partir de données à grands déports. L'approche sélectionnée est une extension de la stéréotomographie, une méthode de tomographie de pente, où les attributs temps de trajet+pente des ondes réfractées et réfléchies sont inversées au sein d'un algorithme hiérarchique multi-échelle. Le potentiel de la méthode est discuté à partir d'un cas synthétique représentatif du champ pétrolier de Valhall, avant l'application aux données réelles du champ de Valhall.

Les derniers chapitres sont consacrés à la reconstruction de plusieurs classes de paramètres dans les approximations visco-acoustiques et visco-élastiques. Afin de réduire la non-linéarité du problème inverse, je propose une méthodologie fondée sur une reconstruction hiérarchique des classes de paramètres et sur l'introduction progressive des différentes composantes de données dans l'inversion. Je teste dans un premier temps différentes stratégies d'inversion sur le cas synthétique de Valhall pour reconstruire la vitesse de compression, la densité et l'atténuation, avant de les mettre en application sur les données réelles dans l'approximation visco-acoustique. J'aborde dans un deuxième temps la reconstruction de la vitesse de cisaillement à partir des trois composantes des capteurs (hydrophone + géophones vertical et horizontal). La qualité des résultats est évaluée par différents outils et une interprétation géologique des résultats est proposée.

Abstract

Geophysics is a very insightful field to gain inferences on the internal structure of the Earth at different scales, and on the past events related to its slow evolution. It also constitutes a major issue in our society, as the resources are becoming increasingly rare, and because it is of great consideration for town and country planning. Full waveform inversion is an optimization process differs from other seismic imaging approaches by its ability to extract the full information content of the seismic data to construct a high-resolution quantitative imaging of one or more physical parameters. I discuss in my thesis several issues related to this method and I illustrate them with applications to multi-component Ocean Bottom Cable (OBC) data recorded at the Valhall oilfield in North Sea.

I first discuss the footprint of anisotropy in vertical transverse isotropic media on the seismic Valhall dataset by comparing the subsurface models built by full waveform inversion when anisotropy is taken into account or not in the seismic modeling. I show some bias in the velocity reconstruction when the imaging is performed in the isotropic approximation, these biases leading to mispositioning of the reflectors at depth.

I then address the problem of constructing an initial velocity model for full waveform inversion from wide-aperture data. The selected approach is an extension of stereotomography, a slope tomographic method, where the traveltimes and the slopes of the reflected and refracted arrivals are sequentially introduced in the inversion process through a multiscale approach. The potential of the method is discussed based on the synthetic Valhall case study before the application to the real Valhall data-set.

The final chapters are devoted to the reconstruction of several classes of parameters within the visco-acoustic and visco-elastic approximations. In order to reduce the nonlinearity of the inverse problem, I propose a methodology based on the hierarchical reconstruction of the parameter classes, and on the progressive introduction of the different data components in the inversion process. I first test different inversion strategies on the synthetic Valhall case to build the compressive velocity, the density and the attenuation parameters, before their application to the real data within the visco-acoustic approximation. I reconstruct in a second step the shear velocity from the three components of the sensor (one hydrophone and a vertical and horizontal geophones). Quality of the results is assessed with various tools, and a geological interpretation of the results is proposed.

Acronymes

4C : données quatre composantes (1 hydrophone et 3 géophones)

CIG : *Common Image Gathers*

DF : Différences Finies

EF : Éléments finis

FWI : *Full Waveform Inversion*, soit *Inversion des formes d'ondes*

GD : méthode Galerkin Discontinue

Go : Giga octets

I_P : impédance de compression

I_S : impédance de cisaillement

P_k : polynôme de degré k

Q_P : facteur de qualité de la vitesse de compression (l'inverse de l'atténuation)

Q_S : facteur de qualité de la vitesse de cisaillement

ρ : masse volumique

OBC : *Ocean Bottom Cable*

PML : *Perfectly Matched Layer*

ν : coefficient de Poisson

RTM : *Reverse Time Migration*

V_H : vitesse horizontale

V_{NMO} : vitesse *Normal Move Out*

VSP : *Vertical Seismic Profile*

V_{P0} : vitesse verticale de compression

V_P : vitesse de compression isotrope

V_S : vitesse de cisaillement

V_X , V_Z , P (ou T_1) : composantes de vitesse horizontale et verticale, et composante de pression des données.

VTI : *Vertical Transverse Isotropy*

*“ La théorie c’est quand on sait tout et que rien ne fonctionne.
La pratique c’est quand tout fonctionne et qu’on ne sait pas pourquoi.
Ici, nous avons réuni théorie et pratique : rien ne fonctionne... et on ne sait pas pourquoi. ”*

Albert Einstein

*“ Si le problème a une solution il ne sert à rien de s’inquiéter,
mais s’il n’y a pas de solution, s’inquiéter ne changera rien ”*

Mantra bouddhiste

Table des matières

Introduction générale	15
0.1 Enjeux et contexte scientifique	15
0.1.1 Motivations	15
0.1.2 Observations sismiques	16
0.1.3 Contexte de ma thèse	17
0.2 L'imagerie sismique	19
0.2.1 Migration qualitative et quantitative	20
0.2.2 Construction d'un macromodèle de vitesse	22
0.2.3 Inversion des formes d'ondes	27
0.2.3.1 Principe	27
0.2.3.2 Etat de l'art	28
0.2.3.3 Inversion multiparamètre	29
0.3 Plan de la thèse	31
1 Méthodes	33
1.1 Modélisation des ondes sismiques dans les milieux 2D visco-élastiques	33
1.1.1 Équations élasto-dynamiques en domaine fréquentiel	33
1.1.1.1 Cas isotrope	33
1.1.1.2 Cas de l'anisotropie VTI	36
1.1.2 Problèmes direct et inverse	37
1.1.2.1 Principe général	37
1.1.2.2 Factorisation LU par solveur direct	38
1.1.3 Discrétisation de la matrice d'impédance \mathbf{A}	38
1.1.3.1 Introduction	38
1.1.3.2 Différences finies	39
1.1.3.3 Galerkin Discontinu	43
1.2 Problème inverse	47
1.2.1 Introduction	47
1.2.2 Théorie des méthodes locales d'optimisation	48
1.2.2.1 Equations normales des approches linéaires et non-linéaires	48
1.2.2.2 Algorithmes d'optimisation locale	52
1.2.2.3 Calcul du gradient d'une fonctionnelle par la méthode de l'état adjoint	54
1.2.3 Application à l'inversion des formes d'ondes	56
1.2.3.1 Approche fréquentielle	56

TABLE DES MATIÈRES

1.2.3.2	Construction du gradient par la méthode de l'état adjoint, et interprétation du gradient	56
1.2.3.3	Construction et interprétation du Hessien	60
1.2.3.4	Analyse de résolution	61
1.2.3.5	Approches multi-échelles et algorithme	63
1.2.3.6	Régularisations	65
2	Le champ pétrolier de Valhall	67
2.1	Introduction	67
2.2	Contexte géologique et pétrolier	69
2.3	Etat de l'art des études géophysiques effectuées sur Valhall	73
2.4	Présentation du jeu de données OBC	78
2.4.1	Anatomie des données	78
2.4.2	Composante de pression calculée pour le cas synthétique de Valhall . . .	80
2.4.3	Données 4C du dispositif LoFS	80
3	Inversion des formes d'ondes acoustique monoparamètre	89
3.1	Introduction au chapitre	89
3.2	Empreinte de l'anisotropie du milieu sur une inversion acoustique isotrope . .	90
3.2.1	Présentation de l'article	90
3.2.2	Résumé	91
3.2.3	Introduction	92
3.2.4	Méthodes	95
3.2.5	Application au champ de Valhall	97
3.2.5.1	Contexte géologique et géométrie d'acquisition	97
3.2.5.2	Anatomie de données et évaluation de la qualité des modèles initiaux	98
3.2.5.3	Prétraitement pour l'inversion des formes d'ondes et protocole expérimental	101
3.2.5.4	Résultats d'inversion	104
3.2.5.5	Evaluation de la qualité des modèles	108
3.2.5.6	Discussion	118
3.2.5.7	Conclusion	120
3.2.6	Résultats complémentaires à l'article	121
3.2.7	Conclusion de l'article	124
3.3	Analyse paramétrique de l'inversion acoustique	126
3.3.1	Non-linéarités introduites par la surface libre - apport du préconditionnement - sensibilité à l'offset	126
3.3.2	Sensibilité aux modèles initiaux anisotropes, à la fréquence initiale et au préconditionnement des données, relativement à une inversion isotrope .	127
3.3.3	Sensibilité de l'inversion aux imprécisions de la partie superficielle du modèle initial	133
3.3.4	Sensibilité à la méthode d'optimisation et à la régularisation	133
3.3.5	Sensibilité aux normes sur données réelles	136
3.4	Conclusion du chapitre	137
4	La stéréotomographie comme outil de construction d'un modèle initial pour	

l'inversion des formes d'ondes	139
4.0.1 Présentation de l'article	139
4.0.2 Résumé	140
4.0.3 Introduction	140
4.0.4 Méthodes	143
4.0.4.1 Stéréotomographie	143
4.0.4.2 Inversion des formes d'ondes dans le domaine fréquentiel . . .	146
4.0.5 Cas synthétique de Valhall	147
4.0.5.1 Protocole expérimental	147
4.0.5.2 Inversion des courts-offsets	151
4.0.5.3 Inversion de 16 km d'offset	151
4.0.5.4 Inversion de 24 km d'offset	154
4.0.6 Application au champ de Valhall	159
4.0.6.1 Inversion stéréotomographique	161
4.0.6.2 Inversion des formes d'ondes	162
4.0.7 Conclusion	166
4.0.8 Appendice A : Régularisation d'une inversion aux classes multiples de données et de paramètres	172
4.0.9 Conclusion de l'article	173
5 Inversion visco-acoustique des formes d'ondes - Reconstruction de la vitesse P, de la densité ou l'impédance, et de l'atténuation	175
5.1 Introduction à l'article	175
5.1.1 La vitesse des ondes P	176
5.1.2 Masse volumique et impédance	176
5.1.3 L'atténuation	177
5.1.4 Position du problème multiparamètre	180
5.2 Résumé	182
5.3 Introduction	183
5.4 Inversion des formes d'ondes visco-acoustique	184
5.5 Reconstruction de la densité ou de l'impédance	187
5.5.1 Paramétrisations et diagrammes de rayonnement	187
5.5.2 Cas synthétique de Valhall	190
5.5.2.1 Modèles initiaux, acquisition et protocole d'inversion	190
5.5.2.2 Résultats d'inversion : reconstruction simultanée avec les paramétrisations (V_P , ρ) et (V_P , I_P)	191
5.5.2.3 Résultats d'inversion : reconstruction hiérarchique avec les paramétrisations (V_P , ρ) et (V_P , I_P)	194
5.6 Reconstruction de la vitesse, de la densité et du facteur de qualité Q_P	197
5.6.1 Cas synthétique de Valhall	198
5.6.1.1 Sensibilité des données et de l'inversion monoparamètre aux modèles initiaux de densité et d'atténuation	198
5.6.1.2 Stratégies d'inversion	200
5.6.1.3 Résultats d'inversion	202
5.6.2 Application aux données OBC du champ de Valhall	207
5.6.2.1 Modèles initiaux et protocole d'inversion	208
5.6.2.2 Résultats d'inversion et évaluation de la qualité des modèles . .	209

TABLE DES MATIÈRES

5.7	Conclusion	218
5.8	Conclusion à l'article	218
6	Inversion visco-élastique des formes d'ondes - Reconstruction des vitesses P et S	221
6.1	Introduction à l'article	222
6.2	Résumé	222
6.3	Introduction	223
6.4	Inversion des formes d'ondes visco-élastique de données multi-composantes . . .	225
6.5	Stratégies multi-échelles et paramétrisation de l'inversion des formes d'ondes .	228
6.5.1	Approches hiérarchiques pour une inversion multiparamètre	228
6.5.2	Choix de la paramétrisation et ses motivations	228
6.5.3	Diagrammes de rayonnement des paramètres V_P et V_S	229
6.5.4	Modèles initiaux et stratégies d'inversion dans cette étude	231
6.6	Protocole pour les données et l'inversion des formes d'ondes	232
6.7	Résultats d'inversion	234
6.8	Evaluation de la qualité des modèles	234
6.8.1	Comparaison avec des logs soniques	234
6.8.2	Accord avec les données	236
6.8.2.1	Anatomie des données	236
6.8.2.2	Modélisation sismique	238
6.8.3	Estimation des ondelettes source	239
6.9	Dérivation d'autres quantités physiques et interprétations géologiques	248
6.9.1	Structure superficielle du modèle de vitesses d'ondes S	248
6.9.2	Coefficient de Poisson et autres combinaisons de V_P et V_S	248
6.9.2.1	Réflecteurs à 0.6 km, 1.25 km, 1.5 km et 2.4 km de profondeur	252
6.9.2.2	Au niveau du réservoir	252
6.9.3	Paramètres I_P, I_S, Q_S	253
6.9.4	Indicateurs d'hydrocarbures	253
6.10	Conclusion	255
6.11	Conclusion aux chapitres multiparamètres	256
7	Conclusions et perspectives	257
7.1	Un travail s'inscrivant dans la prise en compte des dispositifs grands-angles . .	257
7.2	Implications de mon travail	258
7.3	Perspectives	259
	Bibliographie	262
A	Implémentation de la paramétrisation (V_P, I_P, V_S)	287
B	Synthèse de la mémoire prise et du temps consommé pour les différentes classes de tests réalisées	291

Introduction générale

Sommaire

0.1	Enjeux et contexte scientifique	15
0.1.1	Motivations	15
0.1.2	Observations sismiques	16
0.1.3	Contexte de ma thèse	17
0.2	L'imagerie sismique	19
0.2.1	Migration qualitative et quantitative	20
0.2.2	Construction d'un macromodèle de vitesse	22
0.2.3	Inversion des formes d'ondes	27
0.3	Plan de la thèse	31

0.1 Enjeux et contexte scientifique

0.1.1 Motivations

De la même manière qu'en astronomie où l'observation des confins de l'Univers permet de remonter le temps, l'étude de l'intérieur de la Terre nous renseigne sur les phénomènes passés liés à la lente évolution de la Terre. L'astronomie est considérée comme la plus ancienne des sciences, les anciens pouvant prédire la mécanique céleste à partir d'observations directes du ciel. À l'inverse, la compréhension des processus internes n'est que très récente, les éléments collectés en surface n'ayant présenté un intérêt qu'à partir du 17^{ième} siècle¹. Ces enregistrements de surface sont pourtant riches en information sur les phénomènes passés, mais il est très difficile de les décoder car altérés par leurs transports au sein de la Terre (cellules de convection, subductions...). De plus, nous ne pouvons pénétrer la Terre que sur une très faible épaisseur, le plus profond forage n'atteignant *que* 12262 mètres².

La géophysique a pour vocation de sonder la Terre par des techniques de mesures indirectes, et permet de comprendre et voir ce à quoi nous n'avons pas accès. Les disciplines de la géophysique telles que la gravimétrie, le géomagnétisme, la sismologie, les méthodes électriques ou électromagnétiques (pour la très proche surface) parmi d'autres, se distinguent par le type d'observable et par le phénomène physique à son origine. La sismologie est sans doute la discipline la plus riche en applications et au meilleur pouvoir de résolution. La Terre n'est ni opaque, ni transparente aux ondes mais translucide. Cela signifie que les ondes sismiques, en

1. publication de l'ouvrage *Théorie de la Terre* en 1644 par René Descartes

2. forage de Zapoliarny en Russie effectué de 1970 à 1989

traversant la Terre, rapportent des signatures de ses propriétés physiques à la surface où il nous faut les enregistrer. En sismologie *passive*, les tremblements de terre sont utilisés comme source sismique afin d'imager la structure interne de la Terre (Oldham (1906); Gutenberg (1914) et Lehmann (1936) pour les premiers succès), mieux comprendre sa *dynamique*, les zones de failles ou pour localiser et (tenter de) prévoir les séismes et les tsunamis. Classiquement, les temps d'arrivée des ondes sont utilisés pour obtenir une image en terme de vitesse de propagation. Les oscillations libres de la Terre, excitée lors d'un fort tremblement de terre, peuvent aussi être utilisées pour renseigner sur sa structure interne (Dahlen & Tromp, 1998). En sismologie active ou sismique, les sources sont générées artificiellement pour étudier le sous-sol, de la proche surface jusqu'à potentiellement la croûte terrestre, voir la lithosphère. Elles ont connu un développement foudroyant en raison des enjeux sociétaux majeurs associés comme la recherche des sources d'énergie et des ressources minérales mais aussi pour mieux gérer le territoire. Cette discipline a pris son plein essor à partir des années 1920, où on commence à l'utiliser pour détecter des ressources d'hydrocarbures et des minerais. De nombreuses autres motivations sont apparues, en relation avec la détection d'aquifères ou celles liées à une meilleure prise en compte du risque dans l'aménagement du territoire. Nous pouvons citer la détection de cavités souterraines pouvant s'effondrer (Gélis, 2005), la détection de poches d'eau au sein des glaciers pouvant brutalement se libérer (Descloitres *et al.*, 2008), l'étude de la solidité d'un édifice tel qu'un barrage, la détection des zones d'instabilités gravitaires (Bichler *et al.*, 2004; Heincke *et al.*, 2009) préalablement à la construction de maisons et de routes, ou encore le suivi dans le temps du stockage de CO₂ (Morency *et al.*, 2011) et des déchets radioactifs... Il s'avère ainsi que, en plus de l'enseignement sur l'histoire de la Terre, l'étude du sous-sol constitue un enjeu présent et d'avenir de première importance.

0.1.2 Observations sismiques

Les ondes se propageant dans la Terre sont les ondes directes, les ondes réfractées ou coniques ou réfléchies sur une interface géologique, les multiples (internes ou liés à la réverbération dans un guide d'eau), les ondes diffractées par une hétérogénéité, et les ondes de surface, en présence d'un fort contraste de type air ou eau-solide.

En sismique terrestre, les sources utilisées peuvent être une chute de masse, des explosifs ou, plus généralement, un camion vibreur, qui exerce une pression au sol à intervalles rapides et réguliers via une plaque en acier. Une étape supplémentaire de traitement est alors nécessaire pour compresser le signal long émis en une impulsion brève (par corrélation du signal source avec le signal reçu (Lavergne, 1986)). Les ondes sont enregistrées par des capteurs directionnels appelés géophones, mesurant la vitesse ou l'accélération d'une particule au passage de l'onde, suivant les trois directions spatiales. Ces instruments multicomposantes permettent de distinguer les différents types d'ondes (P, S ou de surface), caractérisés par la polarisation de la particule lorsqu'elle vibre au passage de l'onde.

On distingue la sismique réflexion de la sismique réfraction, ou grand-angle, selon la longueur du dispositif et du type d'ondes enregistrées (réfléchi et/ou réfracté). Dans les deux cas en sismique marine, un canon à air, rechargé par un compresseur embarqué et tracté par un bateau, émet une source explosive à quelques mètres de profondeur. Les ondes émises en atteignant le fond marin sont transmises dans le milieu, qu'elles traversent à plus ou moins grande profondeur avant de revenir en surface pour être enregistrées par une flûte dans l'eau composée d'hydrophones, mesurant la pression. Le dispositif d'enregistrement peut être composé d'une

ou plusieurs flûtes sismiques, selon le type de sismique désiré (réflexion ou réfraction). La mise en oeuvre de plusieurs flûtes tractées par plusieurs bateaux nécessite cependant une logistique plus complexe.

En sismique réflexion, les paramètres de l'acquisition tels que l'intervalle de temps entre deux tirs, la vitesse du bateau, l'espacement entre récepteurs sont calculés de sorte que les traces issues de différents tirs éclairent les mêmes points en profondeur. Le nombre de fois qu'un point est éclairé s'appelle la couverture, et détermine le nombre de traces présent dans un profil en point milieu commun (CMP, pour *Common Mid Point*). En sismique réfraction, l'intervalle entre deux tirs est plus espacé en raison des plus grands déports, et les sources doivent être plus puissantes. La profondeur d'investigation sera fonction du contenu basse fréquence de la source et du déport maximum de l'acquisition. Pour des raisons techniques, les sources actives ne génèrent pas d'ondes très basses fréquences (inférieures à 2 Hz). Nous verrons que cela peut être problématique, car ces basses fréquences limitent le risque de saut de phase lors de l'application de la méthode d'inversion des formes d'ondes, qui procède par inversion des basses vers les hautes fréquences. Des progrès dans cette direction sont cependant réalisés. Soubaras & Whiting (2011) préconise ainsi d'utiliser une profondeur de flûte variant avec l'offset, pour obtenir après le traitement des données un spectre fréquentiel plus large. Une autre catégorie de récepteurs appelés OBS, pour *Ocean Bottom Seismometer*, peut être utilisée pour enregistrer des ondes issues de séismes locaux ou de sources actives (avec un réseau plus dense de stations). Ce sont des capteurs souvent 4 composantes (3 géophones, 1 hydrophone) déposés sur le fond marin pour une durée déterminée (jusqu'à quelques mois en sismique passive). Ils se sont popularisés dans les années 70-80, afin de compenser la couverture très inégale des récepteurs, essentiellement localisés à terre, alors que les océans recouvrent 70% de la surface du globe. On les utilise dans de nombreuses études, pour des localisations de séismes, ou pour imager des structures à différentes échelles d'exploration : d'ordre kilométrique (Jin *et al.*, 2000), crustale (en sismique réfraction) (Operto & Ph. Charvis, 1996; Operto *et al.*, 2006) et asthénosphérique (Shiobara *et al.*, 2010). Par ailleurs, des câbles de fond de mer appelés OBC, pour *Ocean Bottom Cable*, contiennent des récepteurs à 4 composantes et peuvent rester en place plusieurs années. Ils sont particulièrement adaptés aux expériences de sismique active 4D où les campagnes sismiques sont régulièrement répétées, en tirant bénéfice d'un faible coût d'entretien (Kommedal *et al.*, 2004; Barkved *et al.*, 2010a).

0.1.3 Contexte de ma thèse

Pour situer mon travail parmi ce panel très large de méthodes et de motivations, je dois préciser que ma thèse s'est déroulée dans le cadre du consortium pétrolier SEISCOPE, actuellement financé par neuf compagnies pétrolières (voir Figure 1). Mon travail a été dirigé par mes directeurs de thèse Stéphane Operto et Jean Virieux en interaction avec une quinzaine de participants localisés à ISTerre (Grenoble) et à Géoazur (Sophia-Antipolis, Villefranche/Mer). Ce consortium est dédié à l'imagerie sismique haute résolution à partir de données grands-angles par la méthode d'inversion des formes d'ondes complètes, les cibles imagées étant d'échelle crustale à kilométrique. Les étudiants travaillent chacun sur un aspect spécifique de la méthodologie et de ses applications, traitant par exemple de l'approximation 2D ou 3D, de l'anisotropie, des télé-séismes, etc... L'inversion des formes d'onde (soit FWI, pour *Full Waveform Inversion*) est une méthode ayant pour objectif l'extraction aussi complète que possible de l'information contenue dans les données sismiques (en terme d'amplitude et de phase).

SEISCOPE Annual Meeting – 2010



FIGURE 1 – Photo prise lors de l'Annual Meeting 2010 à Villefranche sur Mer, sur laquelle sont présents les représentants des sponsors ainsi qu'une partie des membres, actuels et anciens, de SEISCOPE (dont le nom est annoté).

L'objet d'étude de ma thèse est le champ pétrolier de Valhall situé en mer du Nord, d'échelle kilométrique, pour laquelle j'ai appliqué la méthode d'inversion des formes d'ondes à partir de données OBC 4 composantes (Figure 2). Ce champ pétrolier est exploité par BP³ depuis 1982, et est très bien contraint par de nombreuses études à son sujet (Munns, 1985; Hall & Kendall, 2003; Olofsson *et al.*, 2003; Sirgue *et al.*, 2010; van Gestel *et al.*, 2011), ainsi que par de nombreux puits d'exploration et de production. Pour illustrer l'histoire géologique de la zone d'étude, nous montrons sur la Figure 3 le modèle de vitesse synthétique de Valhall, représentatif de la géologie du champ. Valhall est un ancien graben ayant joué au Trias, et fut repris en compression au crétacé supérieur. Cette phase de compression est à l'origine d'un anticlinal dont le toit forme un piège pour les fluides du réservoir. Une importante épaisseur de sédiments sus-jacents s'est accumulée au Tertiaire dans un contexte de subsidence. Un nuage de gaz est présent dans ces sédiments, et forme un véritable écran sismique affectant sensiblement la qualité des images migrées (O. Barkved, 2003). Par ailleurs, Valhall est un milieu marin peu profond, dans lequel de forts multiples se propagent au sein d'une fine couche d'eau de 70 m d'épaisseur, à l'origine d'une plus grande complexité des données. La modélisation du champ d'onde complet et la prise en compte de l'ensemble des arrivées (courts et grands angles), donne à l'inversion des formes d'ondes le potentiel d'imager la zone avec détail malgré ces difficultés.

3. British Petroleum

Une application aux données réelles constitue aussi une source d'inspiration, permettant de soulever de nouvelles questions et de remédier aux problèmes se présentant. Initialement, le choix de la modélisation à deux dimensions a été motivé par le fait que l'on pouvait tester de nouvelles approches et réaliser des analyses paramétriques relativement rapidement, pouvant s'avérer utiles avant que Vincent Etienne, Guanghui Hu et Romain Brossier n'entament les applications de la méthode à trois dimensions sur la zone. Mon rôle au sein du consortium a été d'être un des principaux acteurs du traitement et de l'interprétation de ce jeu de données, en interaction avec les responsables du champ de la compagnie BP. Avant d'aller plus en avant dans la description de mon travail de thèse, il me faut décrire plus précisément le contexte de la sismique active et resituer cette technique d'imagerie parmi celles traditionnellement employées en géophysique pétrolière.

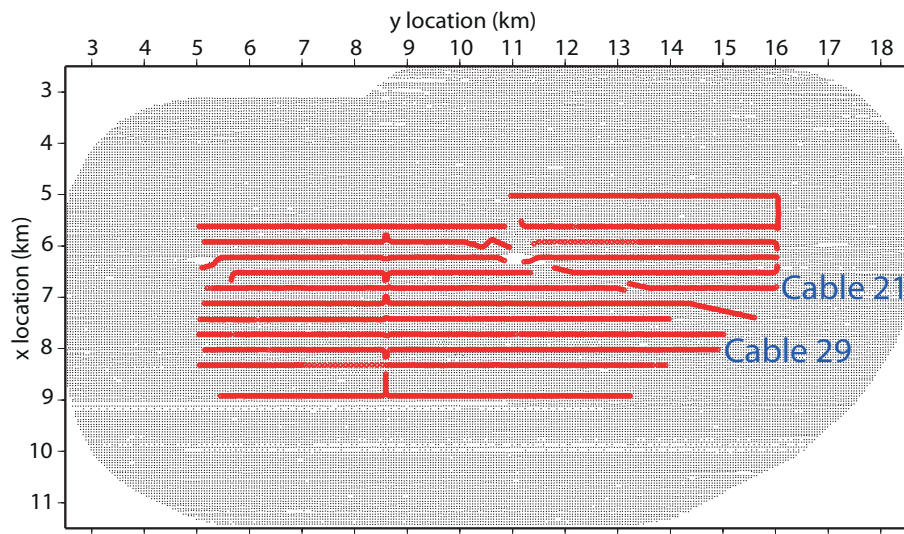


FIGURE 2 – Géométrie de l'acquisition relative dense déployée sur le champ de Valhall en 2003. Les points noirs et rouges représentent la position des tirs (par canon à air) et des câbles de fonds de mer (OBC) 4 composantes, enterrés à 1 m de profondeur (soit 71 m en dessous du niveau marin). L'espacement entre tirs et entre récepteurs est de 50 m, portant le nombre d'OBC et de tirs à environ 2500 et 50000 respectivement. L'acquisition est répétée tous les six mois dans le cadre d'un suivi de production par sismique 4D. La plateforme est localisée à la position ($X = 6\text{km}$, $Y = 11\text{ km}$), à l'origine d'un bruit visible dans les données. Je traite dans ma thèse les câbles 21 et 29 annotés en bleu sur la carte.

0.2 L'imagerie sismique

On distingue deux grandes catégories de méthodes d'imagerie sismique : les méthodes quantitatives, ayant pour but de déterminer la vraie valeur des paramètres physiques telles que les vitesses, afin de mieux qualifier la lithologie ; et les méthodes qualitatives telles que la migration, imageant la réflectivité du milieu. Les données de sismiques réflexion ne sont pas sensibles aux longueurs d'ondes spatiales intermédiaires, en raison du déport limité des acquisitions et du contenu à bande passante limitée du signal source (Jannane *et al.*, 1989). Cela a conduit

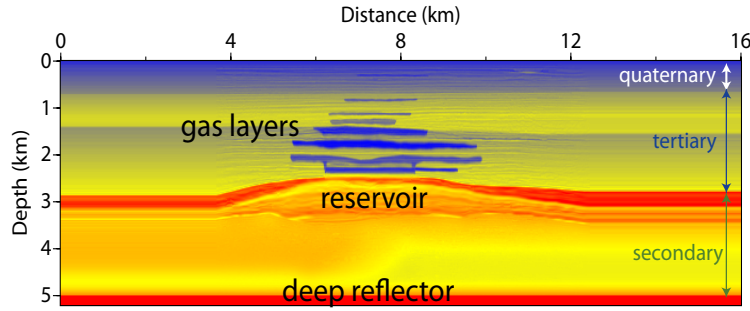


FIGURE 3 – Modèle synthétique de Valhall, représentatif de la géologie du champ de Valhall. Une pile de couches de gaz se trouve dans les sédiments tertiaires. Le réservoir se situe en dessous de l’anticlinal à 2.8 km de profondeur.

à structurer la chaîne de traitement de la sismique réflexion sous forme d’un processus hiérarchique fondé sur un découplage d’échelles (Figure 4) : la construction d’un macromodèle de vitesse lisse, expliquant la cinématique du modèle, est suivie de l’imagerie de la réflectivité par migration. Ce processus est efficace pour des milieux simples mais présente des limitations pour imager des structures complexes telles que des chaînes de chevauchement, des dômes de sel ou des cibles sous du basalte. Bien que l’inversion des formes d’ondes soit en théorie apte à établir la réflectivité, elle est jusqu’à maintenant limitée à la reconstruction des longueurs d’ondes intermédiaires. L’inversion des formes d’ondes a donc été vue comme une alternative possible au processus d’imagerie classique, le modèle de vitesse issu de cette méthode pouvant servir de modèle initial au processus de migration, afin d’en améliorer la qualité en présence de milieux complexes (Operto *et al.*, 2004; Ben-Hadj-Ali *et al.*, 2008).

0.2.1 Migration qualitative et quantitative

La migration est une méthode très utilisée en industrie pétrolière car elle fournit une image structurale haute fréquence du milieu (Biondi, 2006). Elle permet de replacer les réflecteurs à leurs positions exactes et de refocaliser les hyperboles de diffraction à leur apex situé à la position du diffracteur. Par exemple en migration avant-sommation profondeur, un réflecteur penté dans la section non migrée apparaît après migration plus penté et plus court. Cette technique nécessite pour cela un macromodèle de vitesse cinématiquement précis afin de focaliser au mieux les réflecteurs en profondeur (Gray *et al.*, 2001). Il est possible de voir la migration de différentes manières. Une première façon est de considérer qu’un événement dans un sismogramme peut être expliqué par la diffraction d’une onde par un point diffractant. Celui-ci est localisé en profondeur à une position devant être le long d’une isochrone définie par le temps d’arrivée de l’événement et par les vitesses du milieu. Si l’on trace les isochrones de tous les événements d’une trace pour toutes les traces, l’image de réflectivité apparaît par sommation constructive à la position des réflecteurs, et destructive là où il n’y en a pas. Une autre manière a été explicitée par le principe d’imagerie de Claerbout (1970, 1976), où la migration est fondée sur la corrélation du champ incident émis par la source et du champ émis en utilisant les données comme terme de source, rétropropagé en temps depuis les récepteurs. Une corrélation non nulle doit se produire au temps et à la position du réflecteur à l’origine des données observées. Cette intuition des temps coïncidents permet d’avoir des images du sous-sol qualitatives. C’est

Workflow in seismic imaging

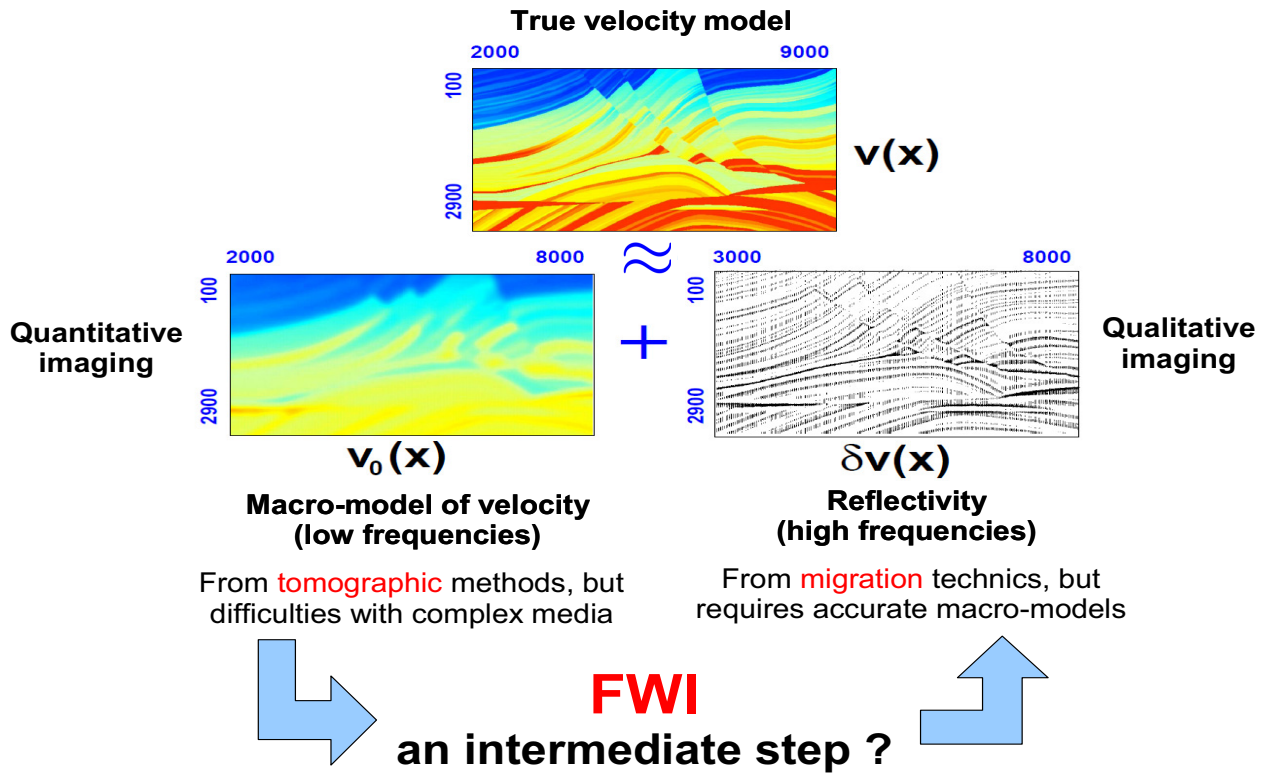


FIGURE 4 – En imagerie sismique, le modèle de vitesse peut être vu en première approximation (signe “ \approx ”) comme la somme d’un macromodèle de vitesse et d’un modèle de réflectivité, renseignant respectivement sur les basses et les hautes fréquences spatiales du modèle. L’inversion des formes d’ondes a le potentiel pour remplir le gap de résolution entre les deux échelles (adapté de Chauris (2000)).

pourquoi Lailly (1983b) et Tarantola (1984a) ont reformulé le principe de migration dans le cadre général de la théorie des problèmes inverses linéarisés, en montrant que la migration s’apparente à la première itération d’un problème inverse qui a pour but la reconstruction des paramètres physiques du modèle. Alors que la migration rétropropage des données pour obtenir une image qualitative, l’inversion des formes d’ondes rétropropage les résidus des données non expliqués par le modèle considéré, pour repositionner les perturbations des propriétés du milieu pouvant potentiellement expliquer ces résidus : on obtient ainsi une image quantitative.

Les migrations s’appuient sur diverses techniques de modélisation, de l’approche asymptotique haute fréquence de la théorie des rais (Červeny *et al.*, 1977), aux méthodes fournissant la solution complète de l’équation d’onde (Gray *et al.*, 2001). On peut par exemple modéliser une acquisition à offset nul grâce au concept de réflecteur explosif, où les sources sont placées au niveau d’un réflecteur et on trace les rais qui partent avec un angle initial perpendiculaire au réflecteur, les réflecteurs étant disposés en surface (Claerbout, 1985). Des arrivées tardives peuvent aussi être considérées grâce aux équations paraboliques *one-way*, modélisant par différences finies une propagation uniquement prograde (Berkhout & Wapenaar, 1989). Les migrations basées sur la résolution des équations d’ondes complètes (*two-way*), appelées *reverse time*

migration, constituent l'outil le plus efficace en migration (Baysal *et al.*, 1983; McMechan, 1989; Chan & McMechan, 1994). C'est la technique se rapprochant le plus de la méthode d'inversion des formes d'ondes.

La linéarisation du problème direct a permis de dépasser le cadre de la migration qualitative et aborder celui de la migration quantitative. Des travaux initiés par Cohen & Bleistein (1979) et poursuivis par Beylkin (1985) et Bleistein (1987) ont cherché à estimer grâce à l'approche asymptotique l'opérateur inverse du problème direct, reliant directement les paramètres du modèle aux données observées. Des méthodes itératives basées sur l'optimisation d'une fonction coût ont été proposées afin de rendre l'approche plus robuste (Jin *et al.*, 1991). Dans l'approche linéarisée rai-Born, les fonctions de Green sont calculées dans un milieu lisse grâce à la théorie des rais, et les perturbations de données et de modèles sont reliées linéairement via l'approximation de Born en diffraction simple. Jin *et al.* (1992) a reformulé les méthodes de migration quantitative dans le cadre théorique des problèmes inverses probabilistes (Tarantola, 1988). Il définit une fonctionnelle dépendant de l'angle d'ouverture au point diffractant, lui permettant d'obtenir un Hessien diagonal. Cette approche est appelée migration/inversion (aux moindres carrés), ou migration avant sommation à amplitudes vraies. Le tracé de rai est réalisé une seule fois dans le modèle initial lisse, car celui-ci est maintenu constant au cours des itérations. Elle fut appliquée en 2D et 3D par Lambaré *et al.* (1992), Thierry *et al.* (1999a,b) et Operto *et al.* (2000). Elle permet d'appréhender diverses acquisitions et divers paramètres (Forgues & Lambaré, 1997) tels que l'atténuation (Ribodetti & Virieux, 1998; Ribodetti *et al.*, 2000a,b). Elles sont cependant limitées par l'hypothèse de diffraction simple de l'approximation de Born.

0.2.2 Construction d'un macromodèle de vitesse

La qualité du macromodèle de vitesse initial nécessaire à l'inversion des formes est primordial en raison du caractère très non-linéaire de cette méthode. L'inversion des formes d'ondes convergera vers un minimum local s'il n'est pas suffisamment précis. Une technique classique est la **tomographie des temps de premières arrivées** (Toomey *et al.*, 1994; Zelt & Barton, 1998). L'accord avec les premières arrivées est critique car ces arrivées sont ensuite inversées par la méthode d'inversion des formes d'ondes, et permettent de reconstruire les grandes longueurs d'ondes du modèle. Cette tomographie n'est cependant pas adaptée en présence de zones à basses vitesses, elle nécessite de grands dépôts pour imager les cibles profondes, et le pointé est souvent interprétatif et fastidieux. Dans cette approche, le modèle de vitesse n'est sensible aux données qu'au niveau du rai, théoriquement d'épaisseur nulle dans le cadre des approches asymptotiques. En **tomographie à fréquence finie** (Luo & Schuster, 1991; Marquering *et al.*, 1999; Dahlen *et al.*, 2000a), le temps d'arrivée d'une phase est fonction de la vitesse dans une zone autour rai, appelée zone de Fresnel. Dans ce cadre, il apparaît que les hétérogénéités au niveau du rai (uniquement) n'influent pas sur la phase de l'arrivée. Zhou *et al.* (2005) par exemple, montre que la prise en compte des noyaux de sensibilité dans une tomographie globale d'ondes de surface amplifie les perturbations de modèle et améliore la résolution de l'image, par rapport au résultat obtenu par approche asymptotique. Le **tomographie des temps de réflexion** (Bishop *et al.*, 1985; Farra & Madariaga, 1988) permet de prédire les arrivées réfléchies, et d'obtenir un modèle à la résolution supérieure à la tomographie en réfraction. Cette approche ne permet cependant pas toujours la prédiction des premières arrivées, elle présente les mêmes difficultés liées au pointé, et souffre de l'ambiguïté vitesse-profondeur à courts offsets. On peut citer l'outil *rayinvr* de Zelt & Smith (1992) permettant d'inverser à la fois les ondes réfléchies, réfractées et tournantes, mais il nécessite a priori fort sur le nombre d'interfaces et leur

épaisseur, difficile à gérer pour les milieux complexes. L'approche semi-automatique de Hobro *et al.* (2002) considère aussi différents types d'arrivées, et permet d'appréhender l'imagerie de structures complexes grâce à l'inversion conjointe d'un modèle de vitesse et de l'épaisseur des couches. Ces deux catégories de paramètres sont définies par des fonctions B-splines pour une représentation lisse du milieu.

Le modèle initial de vitesse utilisé pour la migration profondeur est lui aussi primordial pour permettre une bonne focalisation des réflecteurs. Pour cela, le traitement classique consiste à trier les données en point commun, puis de faire une **analyse de vitesse** afin de trouver les couples temps-vitesse qui vont permettre de redresser les hyperboles de réflexion à tous les temps (Yilmaz, 1987). Cette étape est réalisée via le pointé dans des panneaux de semblance, calculés pour différentes vitesses et à différents instants. Les vitesses obtenues, dites de *sommation* (ou *stacking velocities*), peuvent ensuite être converties en vitesses d'intervalles grâce à la formule de Dix (Dix, 1955). Cependant, ce modèle de vitesse est défini en temps et doit être converti en profondeur, et la validité des vitesses de sommations repose sur l'hypothèse d'un milieu stratifié à faibles variations latérales.

Une migration profondeur avant sommation est considérée comme la plus adaptée pour l'imagerie des milieux complexes à fortes variations latérales et aux pendages non-horizontaux. Par contre, elle est aussi plus sensible au modèle de vitesse que la migration en temps. Cette propriété est d'ailleurs exploitée dans les techniques **d'analyse de vitesse dirigée par la migration** (MVA, pour *Migration Velocity Analysis*) (Yilmaz & Chambers (1984), Chauris (2000)), qui permet de trouver un macromodèle de vitesse dédié à l'optimisation du processus de migration profondeur avant sommation. Contrairement à la tomographie des temps liée à l'optimisation dans le domaine des données, la MVA travaille dans le domaine migré profondeur sur les différentes ondes réfléchies (de type PP, PS, SP, SS). Elle est fondée sur le principe qu'un bon modèle de vitesse doit fournir des images migrées cinématiquement cohérentes, indépendamment du sous-ensemble de données considéré (lié à une classe d'offset ou d'angles). Pour une position X donnée, une coupe est extraite de chaque image migrée associée à un sous-ensemble de données, puis les coupes sont rassemblées sous forme de CIG (*Common Image Gather*, voir Figure 5). La pente du CIG peut être déterminée avec pointé (Liu, 1997) ou sans pointé (Plessix *et al.*, 1995; Symes & Carazzone, 1991). Un CIG plat indique un modèle de vitesse correct, et une pente non nulle permet de mettre à jour le modèle de vitesse de sorte que le CIG dans le nouveau modèle soit plus aplati. Cette méthode itérative, procédant par aller-retour entre le domaine des vitesses et le domaine migré, s'avère être d'un coût numérique important (Woodward *et al.*, 2008), et ne permet pas toujours l'accord aux premières arrivées.

Les méthodes d'inversion dans le domaine de Laplace (Shin & Cha, 2008; Shin & Ha, 2008) permettent de construire un modèle initial lisse sans a priori, en considérant la composante continue des données amorties en temps pour élargir le spectre du signal vers les basses fréquences. Le champ d'onde complet est modélisé en fréquence, en considérant des fréquences purement complexes pour amortir les données. Cette méthode reste difficile à mettre en oeuvre sur données réelles à cause du choix des fréquences complexes, et la forte décroissance de l'amplitude des données avec l'offset fait qu'on a du mal à extraire l'information en présence de bruit. Ce dernier point est partiellement compensé par une fonction coût logarithmique (Shin & Cha, 2008). Cette méthode nécessite néanmoins un pointé des premières arrivées pour muter le signal avant celles-ci. L'inversion de Laplace-Fourier (Shin & Cha, 2009) est équivalente à l'inversion de sismogrammes amortis en temps, en considérant une boucle sur les fréquences réelles et complexes.

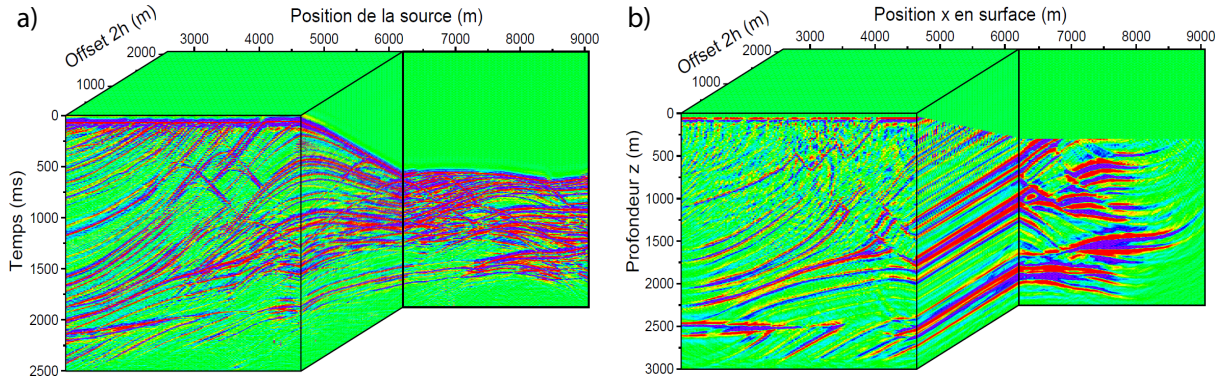


FIGURE 5 – Jeu de données 2D classé en fonction de l’offset (a), et même jeu de données après migration (b). Un CIG est défini pour une position x donnée dans le cube migré. Extrait de Chauris (2000).

Je me suis intéressé à la **stéréotomographie**, une méthode de tomographie de pente fondée sur l’inversion d’événements localement cohérents et des temps d’arrivée de ces événements (Billette & Lambaré, 1998; Lambaré, 2008). Un événement localement cohérent est interprété en terme de réflexion ou de diffraction simple issu d’un point diffractant en profondeur. L’espace des modèles est composé des segments de rais et des paramètres de vitesses (d’ondes P initialement), ces deux catégories de paramètres étant inversées conjointement. Les événements sont pointés simultanément dans le domaine des tirs communs et des récepteurs communs, une seule fois dans tout le processus d’inversion. Les pentes mesurées sont représentatives de l’angle d’incidence du vecteur lenteur en surface. C’est une méthode basée sur les rais : des rais sont tirés depuis chaque point diffractant pour atteindre les sources et les récepteurs, de sorte à pouvoir comparer les pentes observées et calculées; et les dérivées de Fréchet sont calculées par tracé de rais paraxial. Chauris (2000) a établi la connection entre l’analyse de vitesse par migration et les tomographies de pentes. Aplatir localement des événements dans le domaine migré profondeur au niveau de CIGs, est équivalent à faire correspondre les pentes observées et calculées par tracé de rais dans le domaine des temps. Une implication forte est que le pointé peut aussi être fait dans le domaine migré profondeur. Cela s’avère avantageux car les événements sont potentiellement de meilleure qualité dans ce domaine, la migration réduisant les effets de propagation des ondes à l’origine de la complexité sismique dans le domaine temporel. De plus, il est possible dans le domaine migré de travailler directement sur des zones cibles pour affiner une partie du modèle.

Un des points forts de la méthode réside en l’aspect local de la mesure, car il n’est plus nécessaire de suivre une réflexion de manière continue, ni d’interpréter ce qui est pointé. Cela, ajouté au pointé semi-automatique opéré dans des panneaux de semblance (Figure 6), va permettre de considérer un grand nombre de pointés, et donc d’échantillonner de manière complète le milieu par des rais. De plus, la méthode ne souffre pas de limitations liées aux forts contrastes latéraux de vitesse (Billette & Lambaré, 1998), et son modèle initial est un simple gradient de vitesse. L’outil de stéréotomographie a été implémenté pour des milieux en deux et trois dimensions (Chalard *et al.*, 2002), pour des propagations anisotropes (Nag *et al.*, 2006, 2010), et sur divers cas synthétiques et réels (Le Bégat *et al.*, 2000; Billette *et al.*, 2003). La méthode a aussi été étendue à la prise en compte des arrivées converties PS (Alerini *et al.*, 2002; Lambaré & Alérini,

2005; Alerini *et al.*, 2007) pour la reconstruction des vitesses P et S, ainsi qu'à l'utilisation des ondes directes et réfléchies dans un contexte de données de puits (Gosselet & Le Bégat, 2009).

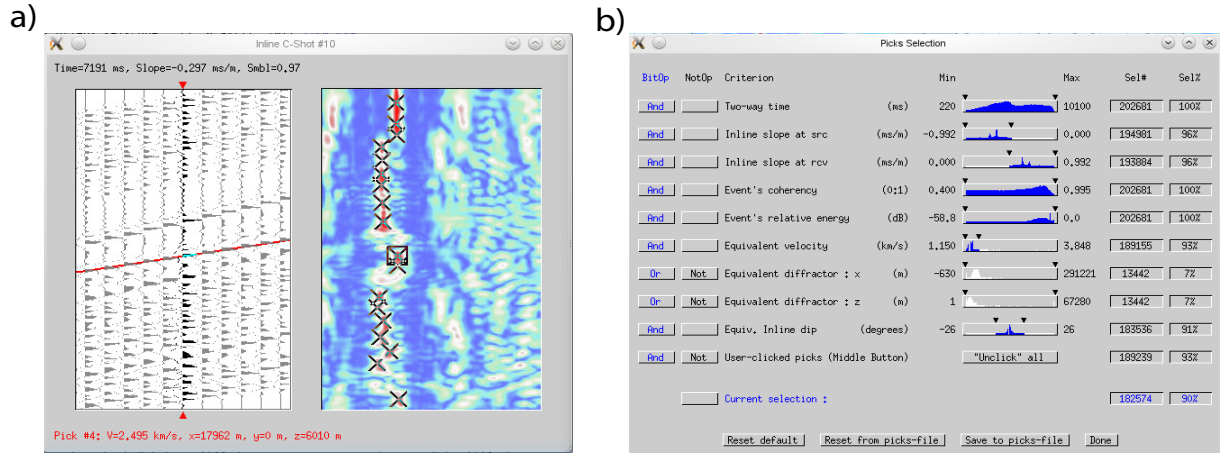


FIGURE 6 – Illustration des deux étapes de pointé en stéréotomographie, réalisées grâce à l'interface graphique développée par Pascal Podvin (de l'École des Mines de Paris), et extraite de notre étude sur le cas synthétique de Valhall. (a) Pointé automatique de l'ensemble des événements localement cohérents au niveau d'une trace centrale. On voit sur le panneau de gauche que cinq traces adjacentes de part et d'autre de la trace centrale sont utilisées pour calculer la semblance. Le panneau de droite représente le panneau de semblance associé aux traces du panneau de gauche, calculé à tous les temps pour toute une gamme de lenteur. Le pointé est réalisé au niveau du maximum de semblance (croix noires). (b) Interface lors de la phase de contrôle qualité du pointé, consistant à omettre les pointés aberrants, apparaissant de manière immédiate grâce à une classification statistique des propriétés des pointés (temps doubles, pentes, énergie du pointé, etc...).

Afin de faire le lien entre les différentes techniques de construction d'un macromodèle, l'inversion des formes d'ondes et la migration, nous montrons sur la Figure 7 les différentes résolutions d'imagerie que les méthodes décrites fournissent sur le cas synthétique de Valhall. Dans cette optique de synthèse, nous montrons aussi sur la Figure 8 la partie réelle du produit de deux fonctions de Green monochromatiques à 5 Hz, correspondant à deux sources situées aux positions du tir et du capteur. Ce produit représente le noyau de sensibilité de l'inversion des formes d'ondes (Woodward, 1992). La zone centrale représente la première zone de Fresnel, dans laquelle tous les diffractants contribuent à construire la première arrivée (comme en tomographie à fréquence finie). Sa largeur L (en vert) donne la résolution obtenue par tomographie des temps de premières arrivées (Williamson, 1991; Pratt *et al.*, 1996). Elle est fonction de la longueur d'onde λ et de l'offset x suivant la loi : $L = \sqrt{\lambda x} = \sqrt{(c x)/f}$, où c et f sont respectivement la vitesse du milieu homogène et la fréquence considérée. Les franges externes sont des isochrones (ou surface équi-phase), c'est-à-dire les surfaces sur lesquelles la réflexion : source S - point diffractant X (partout sur l'isochrone) - récepteur R , met le même temps de parcours. L'épaisseur des isochrones fournit la résolution verticale avec laquelle les réflecteurs sont imagés. L'épaisseur diminuant en s'éloignant du lobe central, on comprend alors le lien existant entre l'angle d'ouverture θ et la résolution de l'imagerie. Ainsi la tomographie en réflexion, associée à des arrivées aux petits angles d'ouverture, sera plus résolutive qu'une tomographie

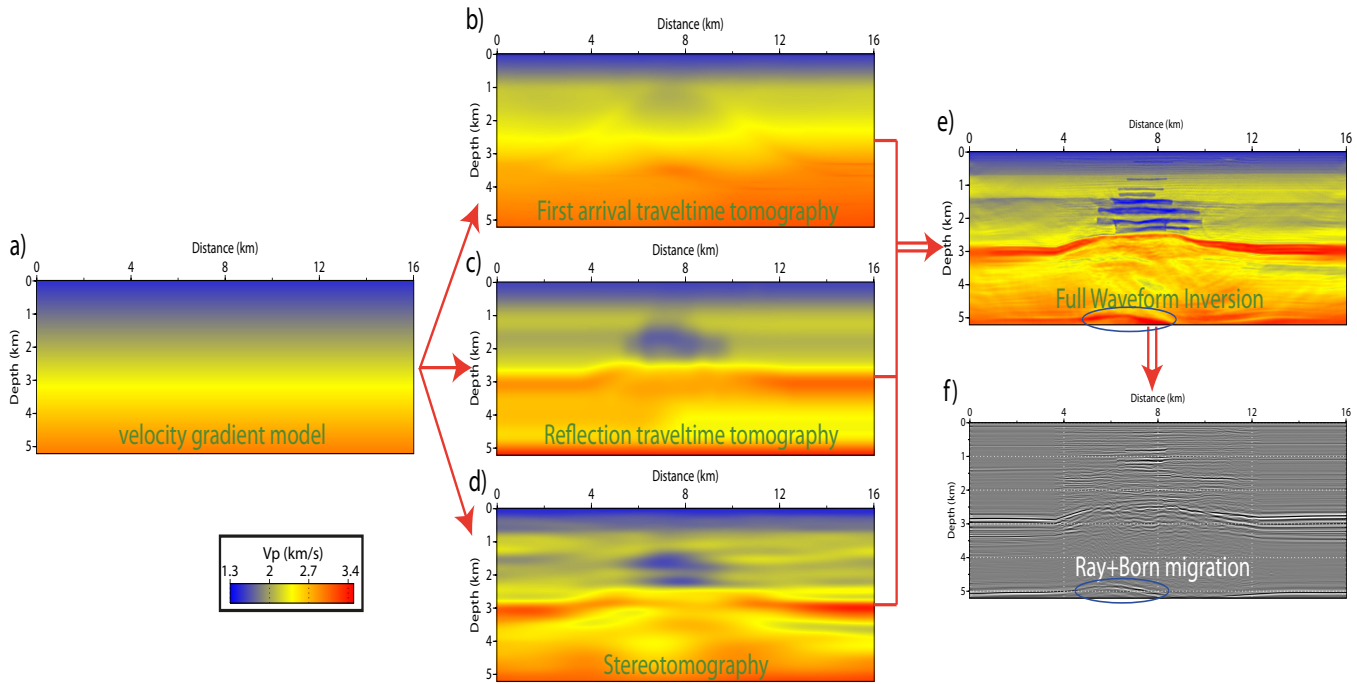


FIGURE 7 – Illustration de la résolution obtenue par différentes méthodes d'imagerie sur le modèle synthétique de Valhall. (a) Modèle de départ à gradient de vitesse utilisé pour la construction d'un macromodèle de vitesse en (b-d). (b) Modèle obtenu par tomographie des temps de premières arrivées en inversant 32 km d'offset. (c) Modèle obtenu par lissage Gaussien du vrai modèle de vitesse (avec une longueur de corrélation de 300 m), représentatif de la résolution que l'on pourrait obtenir (approximativement) par tomographie en réflexion. (d) Modèle obtenu par stéréotomographie en inversant 24 km d'offset et en adoptant la stratégie décrite dans Prioux *et al.* (2012). (e) Modèle obtenu par inversion des formes d'ondes en utilisant le modèle initial en (d). La résolution théorique de la méthode est d'une demi-longueur d'onde (Devaney, 1982; Wu & Toksöz, 1987; Sirgue & Pratt, 2004)). (f) Image de la réflectivité obtenue par migration ray-Born dans le modèle (e) (lissé avec une longueur de corrélation de 75 m). Notez que l'amplitude des réflecteurs au niveau du réservoir (à 2.7 km de profondeur au centre du modèle) est affectée par les piles de gaz sus-jacentes. Notez aussi comment l'imprécision de positionnement du réflecteur profond (ellipse bleue) dans le modèle (e) se répercute dans l'image migrée en (f).

de premières arrivées. Lorsque la fréquence des sources augmente, l'épaisseur des isochrones diminue, synonyme d'un gain en résolution. La stéréotomographie étant basée sur la théorie des rais, elle ne peut pas pointer plusieurs phases au sein de la même zone de Fresnel (pour une trace donnée). Il est donc important que le pointé stéréotomographique soit effectué sur des données hautes fréquences (typiquement entre 10 Hz et 60 Hz) pour lesquelles la largeur des zones de Fresnel est réduite et les arrivées dans les sismogrammes bien séparées. Un plus grand nombre de données stéréotomographiques pourra donc être pointé, conférant ainsi un plus grand pouvoir de résolution à la méthode.

Pour conclure, nous illustrons aussi par cette figure que les méthodes de construction du macromodèle de vitesse n'utilisent que partiellement l'information sismique (premières arrivées,

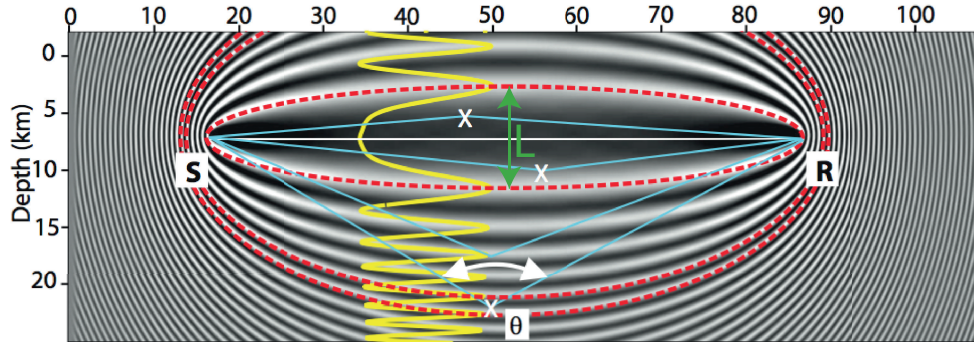


FIGURE 8 – Noyau de sensibilité de l'inversion des formes d'ondes à la fréquence de 5 Hz pour un récepteur localisé à 70 km de la source, dans un modèle homogène à 6 km/s (extrait de Operto *et al.* (2006)). Le lobe central, délimité par les pointillés rouge, représente la première zone de Fresnel et une surface isochrone. Les zones périphériques représentent d'autres isochrones. Les lignes bleues montrent le chemin des diffractions dans la première zone de Fresnel et dans les isochrones. Les oscillations jaunes représentent l'amplitude du noyau de sensibilité.

réflexions, pentes), alors que l'inversion des formes d'ondes a pour objectif d'extraire l'intégralité de l'information sismique.

0.2.3 Inversion des formes d'ondes

0.2.3.1 Principe

L'inversion des formes d'ondes complète est une inversion quantitative basée sur un processus itératif non-linéaire d'optimisation numérique, dédiée à l'obtention d'images hautes résolution d'un ou plusieurs paramètres (Gauthier *et al.*, 1986; Virieux & Operto, 2009). Une inversion globale ou semi-globale ne peut être envisagée pour les milieux 2D ou 3D à cause du coût numérique de ces approches, mais son application au cas 1D a déjà été réalisée grâce à la théorie des rais étendue (Cary & Chapman, 1988; Sambridge & Drijkoningen, 1992). Le champ d'onde doit être calculé à chaque itération dans un nouveau modèle, avec une physique aussi complète que possible (rendant compte des effets 3D, élastiques, anisotropes, de l'atténuation, ...) pour pouvoir reproduire au mieux les données observées : c'est une des significations du terme “*full*” dans FWI. Le problème direct consiste à résoudre le système exact d'équation aux dérivées partielles régissant la propagation des ondes. Il est classiquement résolu par des méthodes de différences finies (Virieux, 1986a; Kolb *et al.*, 1986; Crase *et al.*, 1990), d'éléments finis (Marfurt, 1984; Min *et al.*, 2003; Choi *et al.*, 2008), de volumes finis (Brossier *et al.*, 2008b) ou via des méthodes pseudo-spectrales (Danecek & Seriani, 2008). Une autre raison à l'emploi du terme “*full*” est lié au fait que l'ensemble des arrivées composant le champ d'onde sont utilisées pour mettre le modèle à jour. Cette mise à jour est fondée sur une approche locale d'optimisation où des linéarisations successives de la fonction coût sont réalisées à chaque itération.

0.2.3.2 Etat de l'art

Les bases de la méthode d'inversion des formes d'ondes sont posées par Lailly (1983b) et Tarantola (1984a), avec une formulation en temps. Malgré l'importance des résultats que la méthode semble être en mesure d'obtenir (Gauthier *et al.*, 1986), les applications sur données réelles ne sont apparues que tardivement (Cruse *et al.* (1990); Pica *et al.* (1990); Djikpessé & Tarantola (1999a)). Une des raisons en est le coût prohibitif de la méthode pour les moyens informatiques de l'époque. La méthode constitue aussi un problème inverse très non-linéaire et mal posé, en particulier pour les inversions multiparamètres pour lesquelles un plus grand nombre de degrés de libertés est considéré. Le problème est aussi sous-déterminé en raison des acquisitions de surface, illuminant le milieu de manière incomplète. La méthode s'avère aussi être sensible au bruit polluant les données, à la qualité potentiellement limitée du modèle initial (en particulier pour V_S), au manque de contenu basses fréquences des données, et à la physique des ondes qui repose souvent sur des approximations (2D, acoustique, isotrope) pour des raisons de coûts numériques. Les performances de la méthode ont aussi été fortement limitées par le fait que les applications étaient effectuées pour des dispositifs d'acquisition disposant d'une couverture en offset insuffisante (Mora, 1987, 1988), n'étant sensibles qu'aux hauts nombres spatiaux du modèle (Jannane *et al.*, 1989). L'inversion des formes d'ondes se comportait alors comme une migration, en ne remplissant que la partie haute du spectre des nombres d'ondes. Simultanément à l'arrivée de ressources informatiques plus conséquentes, un tournant, initié par G. Pratt dans les années 1990, repopularisa la méthode en introduisant une formulation fréquentielle du problème direct et inverse (Pratt, 1990; Pratt & Worthington, 1990). Il propose une approche remarquable de cohérence dans l'utilisation des données et des outils numériques (Virieux & Operto, 2009). Il fait la promotion des dispositifs grands-angles tels que des dispositifs d'entre-puits ou de sismique réfraction, qui permettent la reconstruction des longueurs d'ondes intermédiaires et grandes (Pratt *et al.*, 1996; Ravaut *et al.*, 2004). Il montre aussi comment tirer profit de ces géométries pour diminuer le volume de données à inverser. Il a en effet souligné la redondance du contrôle exercé par les fréquences et les angles d'ouvertures sur l'éclairage en nombre d'onde (Devaney, 1982; Wu & Toksöz, 1987). Cette redondance augmente d'autant plus que l'éclairage angulaire est large pour une bande passante donnée (Sirgue, 2003). La redondance de la couverture en nombre d'ondes peut ainsi être réduite en diminuant le nombre de fréquences inversées. Il propose aussi une approche hiérarchique dans le traitement des données permettant de réduire les non-linéarités de l'inversion : les basses fréquences sont inversées avant les hautes fréquences afin de résoudre les bas puis les hauts nombres d'ondes spatiaux du modèle. De plus, il propose un outil numérique, le solveur direct, très efficace (du moins en 2D) pour résoudre l'équation d'ondes en fréquence pour un nombre limité de fréquences et un grand nombre de sources sismiques. Cet outil est fondé sur des méthodes d'élimination de Gauss pour factoriser la matrice d'impédance décrivant le problème direct (MUMPS-team, 2006). Enfin, l'atténuation est facilement implémentée en fréquence grâce à l'introduction de fréquences complexes. Ainsi, les travaux de Pratt ont influencé la manière dont était pensée l'acquisition sismique, en promouvant l'enregistrement des données grands angles et des basses fréquences (Soubaras & Whiting, 2011). Il fut cependant mis en évidence par Sirgue (2003, 2006) et Pratt (2008) que l'inversion des grands offsets s'avère aussi plus sensible aux sauts de phase en raison du plus grand nombre de longueurs d'ondes propagées. De nombreuses applications ont depuis été réalisées en 2D avec ce formalisme (Pratt, 1999; Dessa *et al.*, 2004; Ravaut *et al.*, 2004; Sirgue & Pratt, 2004; Operto *et al.*, 2004; Brenders & Pratt, 2007a).

Grâce aux progrès informatiques, l'inversion des formes d'ondes est rentrée dans une nouvelle ère depuis 2009, date à laquelle l'industrie pétrolière réalisa avec succès les premières inversions de formes d'ondes en 3D sur des données réelles (Sirgue *et al.*, 2009, 2010; Plessix & Perkins, 2010; Plessix *et al.*, 2010b; Vigh *et al.*, 2010; Sirgue *et al.*, 2011). La forte redondance fournie par une acquisition 3D, et l'importante quantité de données grands-angles – grands azimuts, a semble-t-il permis de venir à bout des fortes non-linéarités liées à la méthode et montre d'impressionnants résultats pour imager des structures complexes telles que des diapirs de sel. Ces progrès sont aussi liés à une prise en compte de plus en plus systématique de l'anisotropie dans la modélisation (Woodward *et al.*, 2008), dont nous soulignerons l'importance dans le chapitre 3. Le modèle de vitesse haute résolution est devenu un réel outil pour faire des interprétations géo-structurales du milieu. Ce modèle permet une meilleure planification du forage des puits grâce aux interprétations stratigraphiques que l'on peut en faire, à l'origine d'importantes réductions de coût (Barkved *et al.*, 2010a). L'utilisation de ce modèle en imagerie 4D est aussi à l'origine d'un gain en résolution permettant d'observer des migrations de fluides au sein du réservoir. Les modèles issus de l'inversion des formes d'ondes sont maintenant utilisés comme modèle initial aux techniques de migrations. L'inversion des formes d'ondes est donc devenue un intermédiaire viable entre les macromodèles de vitesse et les images migrées, et fait partie intégrante de la chaîne de production pétrolière. Malgré ces récents succès, les applications restent néanmoins réalisées dans le cadre de l'approximation acoustique monoparamètre.

0.2.3.3 Inversion multiparamètre

L'inversion multiparamètre est un sujet d'un grand intérêt scientifique et plein de promesses. Les paramètres de vitesses de compression et de cisaillement (V_P et V_S), la densité et l'atténuation sont essentiels pour caractériser la lithologie du milieu, et dans le cadre de la surveillance des réservoirs (Wang, 2008) ou du stockage de CO_2 . La connaissance conjointe de V_P et V_S est un indicateur de fluides et de porosité (Domenico, 1984), quantités clés pour la détection de réservoirs (Tatham & Stoffa, 1976). La précision de ces deux paramètres détermine aussi la qualité des corrections statiques, essentielles pour les milieux à forts contrastes en très proche surface (Muyzert, 2000).

Les deux grandes catégories de difficultés des inversions multiparamètres sont liées à l'empreinte variable des paramètres dans les données, en terme de magnitude et de directivité; et aux unités différentes des paramètres, en terme de moyenne et d'écart type. Concernant le premier point, la paramétrisation de l'équation d'onde est cruciale car elle détermine la force avec laquelle les paramètres s'expriment dans les données, et donc la stratégie d'inversion à adopter. Si deux paramètres ont une empreinte de même magnitude dans les données, il est conseillé de les inverser simultanément; alors que si l'un des deux domine l'autre, il est préférable d'adopter une stratégie hiérarchique, en reconstruisant le paramètre dominant avant le paramètre secondaire. D'autre part, la paramétrisation sera non adaptée si elle confère au paramètre inversé une empreinte négligeable dans les données. La paramétrisation détermine aussi le diagramme de rayonnement du paramètre, c'est-à-dire la manière dont un champ diffracté par une hétérogénéité rayonne en fonction de l'angle d'ouverture (Figure 9) (Devaney, 1982; Wu & Toksöz, 1987). Ce rayonnement détermine la résolution avec laquelle un paramètre sera imagé, les grands et les courts angles d'ouverture étant respectivement associés à la reconstruction des grandes et petites longueurs d'ondes spatiales du milieu (Devaney, 1982; Wu & Toksöz, 1987). Ainsi, le paramètre dominant doit avoir un rayonnement aussi large que possible (isotrope) afin de lui garantir une reconstruction haute résolution. Cependant, un compromis

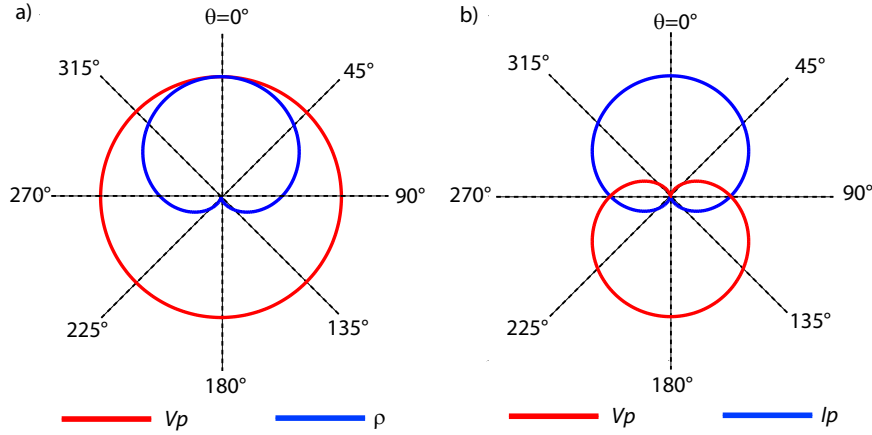


FIGURE 9 – Diagrammes de rayonnement en fonction de l’angle d’ouverture θ , des paramétrisations en vitesse-densité (a) et en vitesse-impédance (b). En (a) par exemple, la courbe rouge indique qu’un champ diffracté par un diffractant de vitesse rayonne de manière isotrope, avec une amplitude plus importante que si l’on considère un diffractant de densité. Selon la paramétrisation, le paramètre de vitesse rayonne de manière isotrope (en a) ou uniquement aux grands angles d’ouverture (en b). Dans la première paramétrisation, un couplage entre la vitesse et la densité se produit pour les courts angles d’ouverture (Forgues & Lambaré, 1997). En d’autres mots, plusieurs combinaisons de paramètres de vitesse et de densité pourront expliquer de la même manière les réflexions à courts offsets. Cependant, le paramètre de vitesse dans la première paramétrisation présente un rayonnement isotrope, lui conférant une reconstruction haute résolution (large bande). Adapté de Virieux & Operto (2009).

doit être trouvé afin que le rayonnement de deux classes de paramètres ne se superpose pas, pour minimiser leur couplage (Figure 9).

L’inversion élastique pour la reconstruction de V_S est particulièrement difficile, expliquant le nombre limité d’applications aux données réelles (Crase *et al.*, 1990; Sun & McMechan, 1992; Igel *et al.*, 1996; Djikpéssé & Tarantola, 1999a; Shipp & Singh, 2002; Sears *et al.*, 2010). Une première raison est la forte non-linéarité du problème inverse, à cause d’un plus grand nombre de longueurs d’ondes propagées pour V_S que pour V_P . De plus, les données terrestres sont affectées par des ondes de surface très énergétiques masquant les arrivées utiles, alors que les données marines ne sont pas assez sensibles aux propriétés élastiques du milieu, en raison de la faible conversion PS se produisant sur le fond marin. Le coût numérique est aussi une limitation importante, la discrétisation devant être beaucoup plus fine pour modéliser la propagation des ondes cisailantes (plus lentes que les ondes de compression). Enfin, comme mentionné précédemment, la plupart des inversions élastiques se limitaient à l’inversion des ondes réfléchies à courts offsets. La non-linéarité de l’inversion élastique fut par la suite réduite en introduisant des données à grands offsets (Freudenreich *et al.* (2001) et Shipp & Singh (2002)), et grâce à une stratégie fondée sur l’inversion de V_S à partir de l’information AVO⁴ portée par les réflexions PP sur les composantes d’hydrophone et de géophone vertical (Sears *et al.*, 2008, 2010).

4. Amplitude Versus Offset

0.3 Plan de la thèse

J'introduis dans un **premier chapitre** le système différentiel qui régit la propagation des ondes et les deux outils de modélisations utilisés au cours de ma thèse pour le résoudre : les différences finies (FD) et les éléments finis discontinus, dit Galerkin Discontinu (DG), développés par Hustedt *et al.* (2004); Sourbier *et al.* (2009a) pour le premier, et Brossier *et al.* (2008b); Brossier (2011a) pour le second. Je présente ensuite le cadre général de la théorie des problèmes inverses linéarisés, ainsi que le développement des équations normales via la méthode de l'état adjoint (Plessix, 2006). J'applique ensuite ces concepts à l'inversion des formes d'ondes complètes et j'interprète le principe d'imagerie qui en découle. **Le second chapitre** fait un état des lieux détaillé des connaissances géologiques et des nombreuses méthodes géophysiques ayant été appliquées sur le champ de Valhall. Cet état de l'art nous servira pour les interprétations géologiques dans les chapitres concernant les imageries multiparamètres. Je présente aussi une interprétation des données 4 composantes et le prétraitement que je leur ai appliqué. **Le chapitre 3** est consacré à l'application de la méthode d'inversion des formes d'ondes monoparamètre au champ de Valhall. La vitesse de compression est alors reconstruite à partir de la composante d'hydrophone dans l'approximation acoustique, isotrope et anisotrope. Je discute en premier lieu de l'empreinte de l'anisotropie VTI dans les données et dans un processus d'imagerie isotrope. Je mets notamment en évidence un biais dans la reconstruction des vitesses et un mauvais positionnement des réflecteurs en profondeur. Je montre dans un deuxième temps le résultat d'analyses de sensibilité à différents paramètres sur les cas synthétique et réel de Valhall. Je discute notamment de l'influence du modèle de départ, de la fréquence initiale, du préconditionnement des données, du lissage des paramètres anisotropes, de l'optimisation et de la régularisation... J'étudie dans **le chapitre 4** la stéréotomographie comme outil de construction d'un modèle de départ à l'inversion des formes d'ondes. Le code de stéréotomographie a été développé au centre de Géophysique de l'École des Mines de Paris. Nous remercions à cet effet Hervé Chauris de nous l'avoir transmis. Traditionnellement restreinte à l'utilisation des données réfléchies à faibles déports, je montre qu'il est possible de combiner les arrivées réfléchies et réfractées grâce à une procédure basée sur l'introduction séquentielle des angles d'ouverture dans l'inversion. L'application de cette stratégie au champ de Valhall montre des résultats prometteurs. Fort de l'expérience acquise par mes tests de sensibilités menés en monoparamètre dans l'approximation visco-acoustique, **les chapitres 5 et 6** traitent de l'inversion des formes d'ondes multiparamètre, dédiée à la reconstruction de la vitesse P, de la densité ou l'impédance, et de l'atténuation dans le chapitre 5, et des vitesses P et S dans le chapitre 6. Les résultats sont présentés sous forme d'articles en préparation. Nous discutons dans ces chapitres à partir des cas synthétique et réel de Valhall, du choix des familles de paramètres et des stratégies d'inversion à adopter, en se basant sur une évaluation quantitative des résultats.

Chapitre 1

Méthodes

Sommaire

1.1	Modélisation des ondes sismiques dans les milieux 2D visco-élastiques	33
1.1.1	Équations élasto-dynamiques en domaine fréquentiel	33
1.1.2	Problèmes direct et inverse	37
1.1.3	Discretisation de la matrice d'impédance \mathbf{A}	38
1.2	Problème inverse	47
1.2.1	Introduction	47
1.2.2	Théorie des méthodes locales d'optimisation	48
1.2.3	Application à l'inversion des formes d'ondes	56

Je présente dans ce chapitre les équations différentielles régissant la propagation des ondes dans les milieux visco-acoustiques et visco-élastiques isotropes et anisotropes VTI, puis les deux principaux outils de résolution de ces équations que j'ai eu l'occasion d'utiliser au cours de ma thèse : les différences finies (DF) et la méthode d'éléments finis de type Galerkin Discontinu (GD). Je rappelle ensuite des notions relatives à la résolution de problèmes inverses via les méthodes locales d'optimisation, dont je me sers pour dériver dans un second temps le formalisme de l'inversion des formes d'ondes. J'explique le rôle et la manière de construire le gradient et le Hessien, puis je décris l'approche multi-résolution employée pour réduire la non-linéarité du problème inverse.

1.1 Modélisation des ondes sismiques dans les milieux 2D visco-élastiques

1.1.1 Équations élasto-dynamiques en domaine fréquentiel

1.1.1.1 Cas isotrope

Origine et formulation

L'inversion du champ d'onde complet exploite la forme des ondes (en terme d'amplitude et de phase), laquelle renseigne sur les propriétés physiques du milieu dans lequel les ondes se sont propagées. La relation entre la forme des ondes et les paramètres du milieu est fournie

par les équations différentielles de l'élastodynamique linéaire. La première relation 1.1 traduit la conservation de la quantité de mouvement (et de masse) d'un élément infinitésimal soumis à une force, telle que celle engendrée par le passage d'une onde. On note par ρ la densité, u_i la composante dans la direction i du vecteur déplacement d'un point élémentaire soumis à des forces extérieures, produites par des densités de forces de volume \mathbf{F} , et des forces de surface fonction du tenseur des contraintes $\boldsymbol{\sigma}$. Le tenseur des déformations $\boldsymbol{\epsilon}$ traduit l'état de déformation local d'un corps déformé. L'équation 1.2 décrit la relation linéaire entre le tenseur des déformations et le déplacement des particules élémentaires, que l'on suppose vérifiée en raison des faibles déplacements engendrés par la propagation d'onde. L'égalité matricielle 1.3 est issue de la loi constitutive de Hooke, et relie linéairement le tenseur de déformation et de contrainte via le tenseur d'ordre 4 des coefficients élastiques du milieu (C_{ijkl}). Cette loi décrit le comportement des milieux 2D isotropes (dans le plan (x-z)), paramétrés par les coefficients élastiques de Lamé λ et μ (le module de cisaillement). Les trois relations décrites s'écrivent

$$\rho \frac{\partial^2 u_i}{\partial t^2} = F_i + \partial_j \sigma_{ij}, \quad (1.1)$$

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad (1.2)$$

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{zz} \\ \sigma_{xz} \end{pmatrix} = \begin{pmatrix} \lambda + 2\mu & \lambda & 0 \\ \lambda & \lambda + 2\mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{zz} \\ 2\epsilon_{xz} \end{pmatrix}. \quad (1.3)$$

L'approximation 2D est utilisée en raison du coût prohibitif des modélisations en 3D, et fait l'hypothèse de l'invariance latérale (suivant y) du milieu et de la source. Le champ d'onde émis par une source ponctuelle et modélisé par un algorithme 2D, génère un champ d'onde identique à ce qu'un algorithme 3D produirait en considérant une source linéique dans le plan perpendiculaire à l'axe de propagation. L'amplitude obtenue par une modélisation 2D décrit donc de manière incorrecte la dispersion géométrique des ondes en 3D, ce qui peut être partiellement corrigé en multipliant par \sqrt{t} l'amplitude des données observées inversées par un algorithme 2D (Bleistein, 1986; Williamson & Pratt, 1995).

La combinaison des trois relations précédentes aboutit au système 1.4, décrivant la propagation isotrope des ondes P-SV polarisées dans le plan (x,z), formulé dans le domaine des fréquences en termes de vitesses de déplacement ($V_i = \partial_t u_i$) et de contraintes (σ_{ij}).

$$\begin{aligned} -i\omega V_x &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right\} + f_x \\ -i\omega V_z &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right\} + f_z \\ -i\omega \sigma_{xx} &= (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_x}{\partial x} + \lambda(\mathbf{x}) \frac{\partial V_z}{\partial z} - i\omega \sigma_{xx0} \\ -i\omega \sigma_{zz} &= \lambda(\mathbf{x}) \frac{\partial V_x}{\partial x} + (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_z}{\partial z} - i\omega \sigma_{zz0} \\ -i\omega \sigma_{xz} &= \mu(\mathbf{x}) \left\{ \frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right\} - i\omega \sigma_{xz0}, \end{aligned} \quad (1.4)$$

où ω représente la pulsation de l'onde, ι le nombre imaginaire (tel que $\iota^2 = -1$), \mathbf{x} le vecteur position, f_i les dérivées temporelles des forces de volume F_i , et σ_{ijo} les termes d'excitation sur les contraintes.

Le système 1.4 est dit d'ordre 1 en raison des dérivées premières d'espace, et s'exprime en fonction de deux champs de vitesse et de trois champs de contraintes. Il peut aussi s'écrire sous la forme d'un système d'ordre 2 ne faisant intervenir que deux champs de vitesse (suivant l'approche parsimonieuse de Luo & Schuster (1990)), en substituant les champs de contraintes du système 1.4 par leur expression fonction de la vitesse (voir équation A.4 dans l'Annexe A). Les outils de modélisation que nous utilisons par la suite sont basés sur un système d'ordre 2 pour le code en différence finie (équation 1.15), et d'ordre 1 pour celui en Galerkin Discontinu. Pour ce dernier, un changement de variable du type $T_1 = (\sigma_{xx} + \sigma_{zz})/2$, $T_2 = (\sigma_{xx} - \sigma_{zz})/2$ et $T_3 = \sigma_{xz}$ est appliqué (Brossier *et al.*, 2008b), afin de mettre le système sous sa forme pseudo-conservative (avec toutes les propriétés physiques du côté du terme de masse du système), tel que :

$$\begin{aligned}
 -i\omega\rho V_x &= s_x \frac{\partial(T_1 + T_2)}{\partial x} + s_z \frac{\partial T_3}{\partial z} + \rho f_x \\
 -i\omega\rho V_z &= s_x \frac{\partial T_3}{\partial x} + s_z \frac{\partial(T_1 - T_2)}{\partial z} + \rho f_z \\
 \frac{-i\omega T_1}{\lambda + \mu} &= s_x \frac{\partial V_x}{\partial x} + s_z \frac{\partial V_z}{\partial z} - \frac{i\omega T_1^0}{\lambda + \mu} \\
 \frac{-i\omega T_2}{\mu} &= s_x \frac{\partial V_x}{\partial x} - s_z \frac{\partial V_z}{\partial z} - \frac{i\omega T_2^0}{\mu} \\
 \frac{-i\omega T_3}{\mu} &= s_x \frac{\partial V_z}{\partial x} + s_z \frac{\partial V_x}{\partial z} - \frac{i\omega T_3^0}{\mu},
 \end{aligned} \tag{1.5}$$

où s_i ($i=x,z$) sont les termes de PML (Perfectly Matched Layer) permettant d'atténuer le champ d'onde dans la couche absorbante (Berenger, 1994). La fonction d'atténuation est définie comme

$$s_i = \frac{1}{1 + i\gamma_i/\omega}, \tag{1.6}$$

où le coefficient γ_i est non nul uniquement dans les couches absorbantes. La formulation pseudo-conservative est plus simple à implémenter dans le cadre des formulations faibles telles que les éléments finis où une intégration sur le volume est faite du système d'équation 1.5. T_1 représente la pression hydrostatique, alors que les termes T_2 et T_3 sont relatifs à la partie déviatorique du tenseur des contraintes, et sont nuls dans l'approximation acoustique isotrope. Dans l'approximation élastique, isotrope et anisotrope, les cinq champs de vitesses et de contraintes sont non nuls.

Différents types d'ondes

Ce système régit la propagation des ondes de compression (ondes P) et de cisaillement (ondes S), de vitesses : $V_P = \sqrt{(\lambda + 2\mu)/\rho}$ et $V_S = \sqrt{\mu/\rho}$ (les plus lentes). Les ondes P et S vibrent respectivement dans la direction parallèle et orthogonale de propagation des ondes. Un autre système d'équation pourrait être déduit des équations 3D de l'élastodynamique, décrivant la propagation des ondes SH polarisées dans les plans (x,y) et (y,z). Le système d'ondes SH est

indépendant du système d'ondes SV uniquement en 2D. En présence d'une surface libre, des ondes de surface de Rayleigh sont générées par la combinaison des ondes P et SV, et des ondes de surface de Love sont issues de la conversion d'ondes SH. Leur amplitude décroît exponentiellement avec la profondeur, mais elles sont très énergétiques (et donc destructrices dans le cas de séismes), et s'atténuent moins vite que les ondes de volume.

1.1.1.2 Cas de l'anisotropie VTI

Système d'équation

Dans les milieux sujets à de l'anisotropie VTI (pour *Vertical Transverse Isotropy*), le nombre d'onde caractérisant la vitesse de compression ($k = \omega/V_P$) est fonction de l'angle de la direction de propagation ($k(\theta)$). Cette anisotropie est relativement commune dans la nature, et se traduit souvent par une vitesse verticale inférieure à la vitesse horizontale, en raison du litage horizontal des sédiments et/ou de l'alignement préférentiel des minéraux dans les argiles. Dans cette approximation, la relation matricielle entre contraintes et déformations s'écrit

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{zz} \\ \sigma_{xz} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{13} & 0 \\ c_{13} & c_{33} & 0 \\ 0 & 0 & c_{44} \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{zz} \\ 2\epsilon_{xz} \end{pmatrix}, \quad (1.7)$$

où les indices numériques des coefficients élastiques sont relatifs à des couples de directions spatiales : $(1, 3, 4) = (xx, zz, xz)$. Deux paramètres supplémentaires sont nécessaires à la description du milieu par rapport au cas isotrope. Le système d'équation correspondant à ce type d'anisotropie peut s'écrire sous la forme

$$\begin{aligned} -\omega\rho V_x &= s_x \frac{\partial(T_1 + T_2)}{\partial x} + s_z \frac{\partial(T_3)}{\partial z} + \rho f_x \\ -\omega\rho V_z &= s_x \frac{\partial T_3}{\partial x} + s_z \frac{\partial(T_1 - T_2)}{\partial z} + \rho f_z \\ -\omega T_1 &= \frac{c_{11} + c_{13}}{2} s_x \frac{\partial V_x}{\partial x} + \frac{c_{13} + c_{33}}{2} s_z \frac{\partial V_z}{\partial z} - \omega T_1^0 \\ -\omega T_2 &= \frac{c_{11} - c_{13}}{2} s_x \frac{\partial V_x}{\partial x} + \frac{c_{13} - c_{33}}{2} s_z \frac{\partial V_z}{\partial z} - \omega T_2^0 \\ -\omega T_3 &= c_{44} s_x \frac{\partial V_z}{\partial x} + c_{44} s_z \frac{\partial V_x}{\partial z} - \omega T_3^0. \end{aligned} \quad (1.8)$$

Ce système régit la propagation d'ondes P, de vitesse verticale $V_{P0} = \sqrt{c_{33}/\rho}$, et d'ondes S de vitesse $V_S = \sqrt{c_{44}/\rho}$. On remarquera que ce système ne peut pas être mis sous une forme pseudo-conservative de manière triviale. Par ailleurs, l'approximation acoustique anisotrope n'est pas une approximation physique, car elle repose sur l'hypothèse que les roches ont un comportement de fluide (pas de cisaillement) intrinsèquement isotrope. Dans cette approximation, le champ T_3 ($= 2 c_{44} \epsilon_{xz}$) est nul, mais contrairement au cas isotrope acoustique, T_2 est différent de 0, sauf si le milieu est elliptique ($\epsilon = \delta$, voir paragraphe suivant).

Signification des paramètres anisotropes

Les paramètres de Thomsen ϵ et δ sont communément utilisés à la place des coefficients élastiques lorsqu'il s'agit d'estimer l'anisotropie du milieu, car ces variables ont une signification

plus physique (Tsvankin, 2001). Le paramètre ϵ est représentatif du rapport entre la vitesse verticale et la vitesse horizontale, alors que δ traduit l'anisotropie subverticale (Thomsen, 1986). Ils s'expriment comme

$$\begin{aligned}\epsilon &= \frac{c_{11} - c_{33}}{2c_{33}}, \\ \delta &= \frac{(c_{13} + c_{44})^2 - (c_{33} - c_{44})^2}{2c_{33}(c_{33} - c_{44})}.\end{aligned}$$

Ces paramètres ont permis à Alkhalifah & Tsvankin (1995) de définir d'autres quantités présentant une relation plus directe entre les paramètres du milieu et leur signature sismique. La vitesse NMO (V_{NMO} pour *Normal Move Out*) et le paramètre η fournissent le cadre pour expliquer la cinématique à courts et grands offsets des temps de trajets des réflexions. Ils sont définis, avec la vitesse horizontale, comme

$$V_{NMO} = V_{P0}\sqrt{1 + 2\delta}, \quad (1.9)$$

$$V_H = V_{P0}\sqrt{1 + 2\epsilon}, \quad (1.10)$$

$$\eta = \frac{\epsilon - \delta}{1 + 2\delta}. \quad (1.11)$$

V_{NMO} explique la cinématique des réflexions à courts offsets, alors que η détermine l'importance de la déviation avec l'offset au comportement hyperbolique des réflexions (Tsvankin, 2001). L'équation non hyperbolique décrivant la cinématique d'une onde réfléchie dans un milieu VTI à une couche prend la forme

$$t^2(x) = t_0^2 + \frac{x^2}{V_{NMO}^2} - \frac{2\eta x^4}{V_{NMO}^2[t_0^2 V_{NMO}^2 + (1 + 2\eta)x^2]}, \quad (1.12)$$

où x représente l'offset, t_0 le temps double de l'onde réfléchie à offset nul. L'introduction de η permet d'avoir une équation à 2 inconnues (V_{NMO} et η) au lieu de 3 (V_{NMO} , δ et ϵ). En sismique réflexion, la reconstruction des paramètres anisotropes se fait suivant une approche hiérarchique, où V_{NMO} est évalué dans un premier temps à partir des réflexions subverticales (deuxième terme de l'équation), puis η est déterminé à partir des réflexions post-critiques dont la cinématique est régit par le troisième terme de l'équation.

1.1.2 Problèmes direct et inverse

1.1.2.1 Principe général

Le système de l'équation d'onde écrit en fréquence peut s'exprimer sous forme matricielle

$$\mathbf{A} \mathbf{p} = \mathbf{s}, \quad (1.13)$$

où \mathbf{A} est une matrice de grande taille et creuse (*sparse*) appelée matrice d'impédance, \mathbf{p} le vecteur donnée formé des composantes de vitesses et/ou de contraintes, et \mathbf{s} le terme de source. Les coefficients de \mathbf{A} sont constitués des paramètres du milieu λ , μ (ou les c_{ij} en anisotrope) et ρ , qui dépendent non linéairement des données \mathbf{p} . Par contre la relation entre les données et la source est linéaire. Le problème direct consiste à calculer le champ d'onde \mathbf{p} aux récepteurs de l'acquisition à partir d'un modèle de paramètres a priori. Le problème inverse a pour vocation de déterminer un modèle de paramètres en minimisant l'écart entre les données calculées et observées.

1.1.2.2 Factorisation LU par solveur direct

La résolution du système d'équation 1.13 en espace/fréquence (1.13) se fait de manière implicite, par résolution d'un système linéaire, par opposition aux schémas explicites en temps où les solutions au temps t doivent être calculées à partir des solutions obtenues aux temps précédents. La résolution ne porte que sur une composante fréquentielle à la fois. L'obtention de sismogrammes en temps nécessite de résoudre le système pour toutes les composantes spectrales du signal, et d'appliquer une transformée de Fourier inverse au résultat obtenu.

Pour résoudre ce système, une factorisation LU de la matrice \mathbf{A} est calculée à chaque fréquence considérée, particulièrement adaptée au format *sparse* de \mathbf{A} . Cette technique consiste en la décomposition de la matrice \mathbf{A} en produit de \mathbf{L} par \mathbf{U} , où \mathbf{L} (pour 'Low') est une matrice triangulaire inférieure, et \mathbf{U} (pour 'Up') une matrice triangulaire supérieure (Duff *et al.*, 1986). Une factorisation LU donne directement la solution exacte, contrairement aux méthodes itératives (telles que GMRES (Saad, 2003)) approximant la vraie solution par itérations successives. L'équation 1.13 devient après factorisation LU : $\mathbf{LU} \mathbf{p} = \mathbf{s}$. L'équation à seconds membres multiples correspondant au problème à plusieurs sources (\mathbf{s}_i), peut s'écrire : $\mathbf{LU} [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n] = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n]$. L'avantage d'utiliser la factorisation LU repose sur le fait que pour un problème à sources multiples, l'opération de factorisation n'est à faire qu'une seule fois (malgré un coût numérique non négligeable), et l'opération pour trouver tous les \mathbf{p}_i se fait très rapidement grâce à un algorithme d'élimination prograde - substitution rétrograde des termes (George & Liu, 1981). Le calcul de factorisation est réalisé grâce à un solveur parallèle dénommé MUMPS, pour : Multifrontal Massively Parallel sparse direct Solver MUMPS-team (2006). D'autre part, une phase d'analyse est effectuée une fois par MUMPS avant la première factorisation LU, et a pour but de réordonner les coefficients de \mathbf{A} afin d'optimiser les opérations de pivots lors de la factorisation et de réduire le remplissage des matrices \mathbf{L} et \mathbf{U} .

La factorisation LU effectuée en parallèle sur différents processeurs peut être particulièrement efficace dans le cadre de calcul hautes performances (approches *multifrontales*, (Duff & Reid, 1983; Amestoy *et al.*, 2000)). Le stockage *distribué* des facteurs LU, plus nombreux que ceux de la matrice \mathbf{A} , sur les différentes unités de calcul (autres que le noeud maître), s'avère indispensable dans le cadre de problèmes de grandes tailles en raison de la mémoire limitée des architectures. Le parallélisme permet de résoudre des problèmes de plus petites tailles sur les différentes unités de calcul. La résolution numérique s'en trouve accélérée, malgré le surcoût en temps, et surtout en mémoire, engendré par les communications supplémentaires.

1.1.3 Discrétisation de la matrice d'impédance \mathbf{A}

1.1.3.1 Introduction

Il existe de nombreuses méthodes de résolution du système d'équation d'ondes présenté précédemment. Dans le cadre de l'inversion des formes d'ondes, la méthode adoptée doit permettre de résoudre ces équations pour un milieu arbitrairement hétérogène, écartant de ce fait les méthodes **analytiques** et semi-analytiques dédiées aux milieux homogènes ou à couches (Bouchon *et al.*, 1989; Garvin, 1956)).

L'approche asymptotique par tracés de rais (Červený *et al.*, 1977) permet de considérer des milieux hétérogènes mais faiblement contrastés (lisses), très utile en raison de son faible coût numérique pour les méthodes de construction de macromodèles telles que la tomographie des temps de première arrivée, la stéréotomographie et pour les approches de migration rai-Born

et *raï-Kirchhoff*. Sous cette approximation haute fréquence, la solution de l'équation d'onde est le produit d'un terme fortement oscillant à hautes fréquences (le terme de phase $i\omega T(x)$, où $T(x)$ est le temps de trajet indépendant de la fréquence) et d'un terme d'amplitude moins oscillant. Pour que l'équation d'onde soit vérifiée après substitution de cette solution, il faut vérifier deux égalités connues sous le nom d'équation Eikonale $([\nabla_x(T(x))]^2 = 1/c^2(x))$, dont la résolution donne accès à une carte des temps de trajet ; et l'équation de transport (de l'énergie) qui permet de calculer l'amplitude des ondes. En général seuls les termes d'ordre '0' de cette dernière équation sont retenus (les plus importants) pour rendre compte de la décroissance géométrique des ondes.

Parmi les méthodes permettant de résoudre l'équation d'onde complète, nous pouvons citer les méthodes **pseudo-spectrales** (Kosloff *et al.* (1990),Fornberg (1996)), pour lesquelles les dérivées spatiales sont calculées dans les domaines de Fourier spatiaux, ce qui permet de transformer les dérivées d'ordre élevés en simples produits. Les dérivées d'ordre élevés ont pour avantage de pouvoir utiliser une grille au pas très lâche, la contrepartie étant l'importante largeur de bande numérique de la matrice d'impédance A associée, peu adaptée au calcul parallèle en raison des communications nécessaires.

Les **éléments finis (EF) continus** (Zienkiewicz *et al.*, 2005) sont une classe de méthodes pour lesquelles la solution est *approximée* dans chaque élément du maillage (segments en 1D, triangles en 2D, tétraèdres en 3D) sur des bases de fonction définies au niveau de *noeuds*. Les éléments finis reposent ainsi sur une formulation dite 'faible' du problème, consistant à vérifier l'équation d'onde en moyenne sur l'élément. Les EF classiques sont dits continus en raison de la continuité de la solution d'un élément à un autre (noeuds communs), comme illustré sur la Figure 1.1a. Cette hypothèse complexifie l'implémentation de cette méthode pour modéliser des discontinuités liquide-solide par exemple (Komatitsch *et al.*, 2000). Les fonctions de bases sont en général polynomiales et sont définies par un degré (ou ordre d'interpolation). La plus grande précision des ordres élevés permet de considérer des éléments de base (ou cellule) de plus grande taille, au détriment d'un plus grand nombre de degrés de libertés. Un des intérêts des EF réside dans la possibilité de mélanger des cellules de différentes tailles adaptées à la vitesse locale (*h-adaptivité*). L'utilisation de triangles ou tétraèdres permet aussi de discrétiser finement des topographies complexes. Le coût numérique de cette méthode reste élevé en raison du nombre important de points par longueur d'onde à respecter pour assurer une précision satisfaisante.

Les **éléments spectraux** (Faccioli *et al.*, 1997; Komatitsch & Vilotte, 1998) sont des EF continus d'ordres élevés, dont les fonctions de bases polynomiales utilisées (de Chebyshev ou de Legendre), différentes de celles des EF classiques, procurent une bonne convergence spectrale (très bonne précision obtenue en montant en ordre). Elles s'appuient sur des maillages non triangulaires (quadrangles en 2D, parallélépipèdes en 3D) peu flexibles.

Nous détaillons à présent deux autres grandes classes de méthodes que nous avons utilisées au cours de la thèse, que sont les **différences finies (DF)** (Virieux, 1984, 1986b) et les **Galerkin discontinus (GD)** (Cockburn *et al.*, 2000).

1.1.3.2 Différences finies

Présentation

L'approche des DF est la plus communément utilisée en raison de sa simplicité théorique et d'implémentation (Virieux, 1986b; Levander, 1988; Moczo *et al.*, 2004; Bohlen & Saenger, 2006).

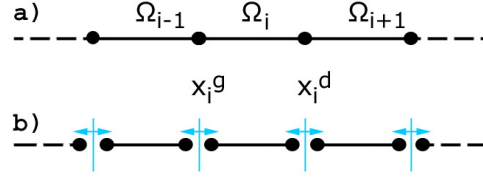


FIGURE 1.1 – (a) Discrétisation par éléments continus P_1 à 2 degrés de liberté sur une corde (cas 1D). Les noeuds frontières sont partagés entre les éléments. (b) Discrétisation par éléments discontinus P_1 . Chaque élément possède son ensemble de noeuds. Les flèches bleues représentent les flux numériques échangés entre chaque élément (extrait de Etienne (2011)).

Contrairement aux EF, les DF reposent sur une formulation forte du problème, où l'équation d'onde doit être vérifiée en chaque noeud d'une grille (traditionnellement) cartésienne. En DF, les dérivées spatiales (et temporelles) du champ d'onde sont approximées par un développement de Taylor à l'ordre plus ou moins élevé selon le schéma de discrétisation choisi. Par exemple un développement à l'ordre 2 estime la dérivée (centrée) selon la direction x de la vitesse V_x au point ' i ' à partir des points voisins $i + 1$ et $i - 1$ comme

$$\frac{\partial V_{X,i}}{\partial x} = \frac{V_{X,i+1} - V_{X,i-1}}{2h}, \quad (1.14)$$

où h est le pas de discrétisation spatial. Il existe différentes façons d'écrire les opérateurs de dérivation dans un repère discret, chacune définissant un schéma de différences finies, sachant qu'un ordre de dérivation plus élevé implique pour l'estimation de la dérivée des points voisins plus éloignés. Les noeuds associés aux vitesses et aux contraintes sont en général décalés d'un pas de $h/2$ (Virieux, 1984), car cette implémentation présente une faible dispersion numérique (voir paragraphe suivant), est indépendante de la valeur du coefficient de Poisson (Moczo *et al.*, 2004), et s'affranchit des problèmes d'interface fluide-solide. Par contre, la grille cartésienne constitue une limitation importante de la méthode DF pour les topographies complexes, à cause des diffractions parasites engendrées si la topographie est discrétisée par des éléments trop grossiers (en marches d'escaliers).

La discrétisation d'une équation différentielle telle que l'équation d'onde sur une grille numérique de pas fini, induit des approximations et donc des erreurs de modélisation matérialisées par de la dispersion et de l'anisotropie numérique. Lorsqu'un champ d'onde monochromatique se propage sur une grille de différences finies, on distingue la vitesse de phase effective du champ sur la grille, de la vitesse de phase théorique (vitesse de propagation du milieu). La différence des deux vitesses constitue la dispersion numérique, fonction de l'échantillonnage du milieu, de la fréquence modélisée et de la précision du schéma numérique utilisé. Par ailleurs, on appelle anisotropie numérique la variation de la vitesse de phase numérique avec la direction de propagation. Afin de minimiser ces effets, un nombre minimum de points doit échantillonner la plus petite longueur d'onde locale, définie comme $\lambda_{min} = c_{min}/f_{max}$, où c_{min} est la vitesse la plus faible présente dans le milieu et f_{max} la fréquence maximum (Marfurt, 1984). À noter que pour les schémas explicites en temps, il faut satisfaire en plus la condition de stabilité de Courant-Friedrichs-Lewy (CFL), dictant que le pas de temps minimum dt_{min} doit être inférieur à $(coeff \cdot h_{min}/c_{max})$, où 'coeff' est un coefficient dont la valeur est fonction du schéma.

Discrétisation et largeur de bande numérique pour les systèmes en fréquence

L'ordre d'un schéma est le nombre de points pris en compte autour du point de colocation dont on veut estimer la valeur du champ, et définit la manière dont est remplie la matrice \mathbf{A} . Par exemple, dans le cas d'un schéma en croix à cinq points (Figure 1.2a), l'écriture discrète d'une ligne de la matrice se ramène à $\mathbf{A}_1 \cdot \mathbf{p}_{i,j-1} + \mathbf{A}_2 \cdot \mathbf{p}_{i-1,j} + \mathbf{A}_3 \cdot \mathbf{p}_{i,j} + \mathbf{A}_4 \cdot \mathbf{p}_{i+1,j} + \mathbf{A}_5 \cdot \mathbf{p}_{i,j+1} = \mathbf{s}_{i,j}$. Il est possible de faire un réarrangement des indices des noeuds i (indice de ligne) et j (indice de colonne) de la grille cartésienne, de la forme : $k = (j - 1) n_z + i$, amenant à des variables V_k ne dépendant plus que d'un seul indice et dont on peut rassembler les valeurs discrètes sous forme de vecteurs de dimension $(n_x \times n_z)$. L'indice 'i' est dit rapide car il augmente d'une unité entre deux points adjacents verticalement. La matrice \mathbf{A} est alors composée de cinq diagonales (Figure 1.2b), l'espace entre les deux diagonales extrémales, dite bande passante numérique, étant de $2n_z + 1$.

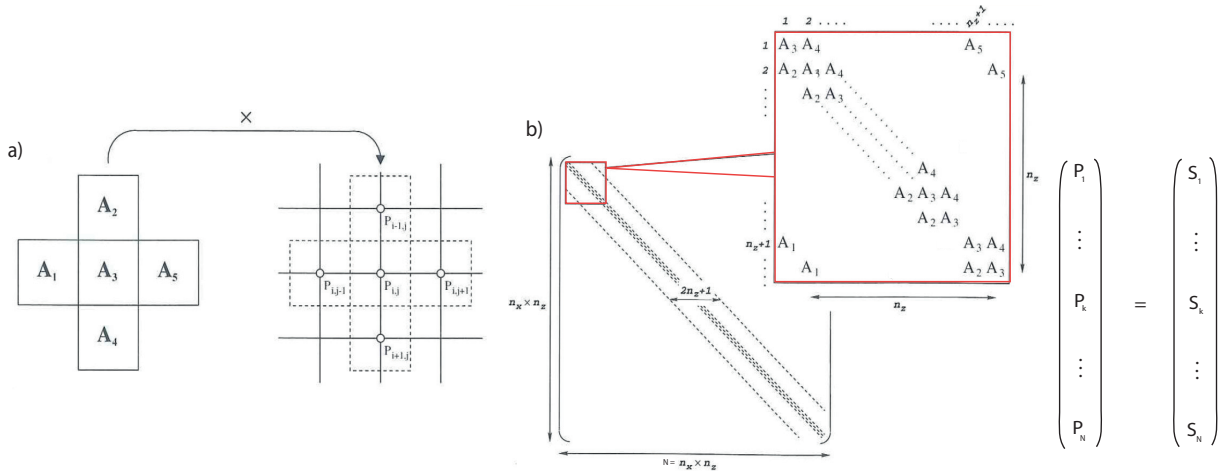


FIGURE 1.2 – (a) Schéma à cinq points en croix. (b) Représentation des termes composant l'équation matricielle : $\mathbf{A} \mathbf{p} = \mathbf{S}$, lorsque le schéma de différences finies utilisé est un schéma à cinq points en croix (adapté de Dessa (2001)).

La mémoire utilisée pour le stockage de la matrice \mathbf{A} factorisée est une considération importante dans les problèmes de différences finies en fréquence. Afin de limiter la mémoire prise, il est réalisé un réordonnancement des valeurs de la matrice (lors de la phase d'analyse de MUMPS dans notre cas), permettant de limiter le remplissage de la matrice après factorisation. Dans le cas d'une grille numérique carrée ($n_x = n_z = n$), la mémoire prise par une matrice issue d'une factorisation LU dont la matrice de départ a subi un réordonnancement est de l'ordre de $n^2 \cdot \log_2(n)$, contre une mémoire de n^3 pour une matrice factorisée directement (George & Liu, 1981). Il a été montré qu'à un coefficient $a_{i,j}$ (de la matrice \mathbf{A} à factoriser) non nul correspond un coefficient $l_{i,j}$ ou $u_{i,j}$ de la matrice \mathbf{L} ou \mathbf{U} non nul; et qu'à un coefficient $a_{i,j}$ nul peut correspondre une valeur non nulle $l_{i,j}$ ou $u_{i,j}$, et ce uniquement si le coefficient $a_{i,j}$ se situe entre les deux diagonales extrémales de \mathbf{A} . Ainsi, afin de limiter le nombre de coefficients non nuls des matrices \mathbf{L} et \mathbf{U} , il importe que la bande passante numérique soit minimisée, favorisant ainsi le schéma de différences finies d'ordre 2 par rapport à celui d'ordre 4. On comprend de même pourquoi l'indice rapide est i , car en général $n_z < n_x$, la largeur de bande valant $(2n_z + 1)$.

Schéma utilisé

Nous avons utilisé le code de Hustedt *et al.* (2004); Sourbier *et al.* (2009a,b), dans lequel est implémentée l'équation d'onde visco-acoustique d'ordre 2 en pression, écrite en fréquence sous la forme

$$\frac{\omega^2}{\kappa(\mathbf{x})}P(\mathbf{x},\omega) + \nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \nabla P(\mathbf{x},\omega) \right) = -S(\mathbf{x},\omega), \quad (1.15)$$

où $\kappa = \rho c^2$ est le module de rigidité (avec c la vitesse), P la pression et S la source. Pour les raisons de mémoire précédemment évoquées, le schéma d'ordre 2 est préféré au schéma d'ordre 4. Afin de compenser la moindre précision associée, le schéma adopté est à grilles mélangées (*mixed-grid*), consistant à combiner l'écriture discrète des termes de différenciation spatiale suivant le repère cartésien et suivant un repère tourné de 45° (Jo *et al.*, 1996; Hustedt *et al.*, 2004). Pour cela, le système d'ordre 1 en espace-fréquence est discrétisé avec des dérivées de précision d'ordre 2 suivant le repère cartésien et le repère tourné. Puis l'approche parsimonieuse de (Luo & Schuster, 1990) est utilisée pour éliminer les vitesses. Les deux opérateurs de discrétisations obtenus sont ensuite combinés, afin d'obtenir une précision d'ordre 4. Le second terme du membre gauche de l'équation 1.15 est ainsi approché par une combinaison linéaire du type : $a \Gamma + (1 - a) \Theta$, où :

$$\Gamma = \left[\frac{\partial}{\partial x} \left(\frac{1}{\rho(\mathbf{x})} \frac{\partial P(\mathbf{x},\omega)}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{1}{\rho(\mathbf{x})} \frac{\partial P(\mathbf{x},\omega)}{\partial z} \right) \right]_{00} \text{ et } \Theta = \left[\frac{\partial}{\partial x'} \left(\frac{1}{\rho(\mathbf{x})} \frac{\partial P(\mathbf{x},\omega)}{\partial x'} \right) + \frac{\partial}{\partial z'} \left(\frac{1}{\rho(\mathbf{x})} \frac{\partial P(\mathbf{x},\omega)}{\partial z'} \right) \right]_{45^\circ},$$

avec (x,z) les coordonnées d'un point dans le repère cartésien, et (x',z') ses coordonnées dans le repère tourné de 45° . Ce schéma de grilles mélangées ainsi que la structure de la matrice d'impédance sont représentés sur la Figure 1.3.

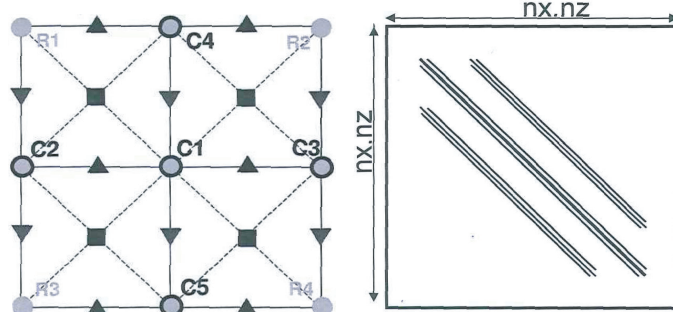


FIGURE 1.3 – (a) Schéma de différences finies dit à grilles mélangées à 9 points (Jo *et al.*, 1996; Hustedt *et al.*, 2004). b) Structure à 9 diagonales de la matrice \mathbf{A} dans ce schéma à grille mélangée. Noter que le support spatial du schéma n'excède pas 2 pas de grille. Extrait de Hustedt *et al.* (2004).

Cette méthode diminue l'anisotropie numérique, mais pas la dispersion de la vitesse de phase. Pour y remédier, il est finalement appliqué la méthode d'étalement du terme de masse : $\left[\frac{\omega^2}{\kappa(\mathbf{x})} P \right]_{i,j}$. Au lieu d'évaluer ce terme au point de collocation C_1 (Figure 1.3), il est estimé par une combinaison linéaire tenant compte de sa valeur à son voisinage, soit

$$\frac{\omega^2}{\kappa_{i,j}(x,z)} \bar{P} = \omega^2 \left[\frac{c}{\kappa_{i,j}} P_{i,j} + d \left(\frac{P_{i+1,j}}{\kappa_{i+1,j}} + \frac{P_{i-1,j}}{\kappa_{i-1,j}} + \frac{P_{i,j+1}}{\kappa_{i,j+1}} + \frac{P_{i,j-1}}{\kappa_{i,j-1}} \right) + \frac{1-c-4d}{4} \left(\frac{P_{i+1,j+1}}{\kappa_{i+1,j+1}} + \frac{P_{i-1,j+1}}{\kappa_{i-1,j+1}} + \frac{P_{i+1,j-1}}{\kappa_{i+1,j-1}} + \frac{P_{i-1,j-1}}{\kappa_{i-1,j-1}} \right) \right]. \quad (1.16)$$

Les coefficients a , c et d sont choisis de sorte que la dispersion de la vitesse de phase soit minimisée ($a=0.5461$, $c=0.6248$, $d=0.09381$ et $1-c-4d=0.1297 \cdot 10^{-5}$ d'après Jo *et al.* (1996)).

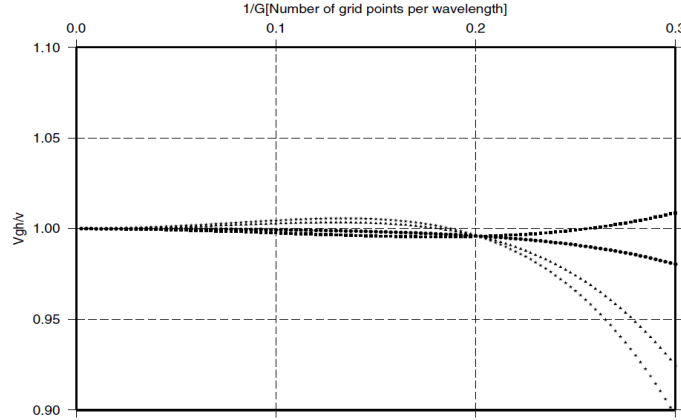


FIGURE 1.4 – Analyse de dispersion de la vitesse de groupe pour un schéma à grilles mélangées (extrait de Hustedt *et al.* (2004)), en fonction de l'inverse du nombre de points par longueur d'onde (G). Les différentes courbes représentent une propagation avec un angle de 0° , 15° , 30° et 45° , représentées avec un figuré en étoiles, triangles, cercles et carrés respectivement.

Finalement l'équation d'onde discrétisée dans le schéma de grilles mélangées avec étalement du terme de masse peut s'écrire : $a \Gamma_{i,j} + (1 - a) \Theta_{i,j} + \frac{\omega^2}{\kappa_{i,j}(x, z)} \bar{P} = S_{i,j}$. Ces stratégies permettent de ramener la discrétisation à quatre points par longueur d'onde, comme illustré sur la Figure 1.4 où nous montrons l'analyse de dispersion réalisée par Hustedt *et al.* (2004) pour ce schéma.

1.1.3.3 Galerkin Discontinu

Présentation

La méthode Galerkin Discontinu (GD) est appliquée dans de nombreux domaines (Cockburn *et al.*, 2004; Bernacki & Piperno, 2003; Jacobs & Hesthaven, 2006), et depuis peu dans celui de la propagation d'ondes sismiques (Ainsworth *et al.*, 2006; de la Puente *et al.*, 2007; Dumbser & Käser, 2006; Käser & Dumbser, 2008; de Basabe *et al.*, 2008; Brossier *et al.*, 2008b; Etienne *et al.*, 2010a). C'est une approche faisant partie de la catégorie des EF discontinus (noeuds non communs, Figure 1.1b), impliquant donc un plus grand nombre de degrés de liberté que les EF continus. Cette approche tire cependant partie d'une adaptation locale des ordres d'interpolation dans les cellules (*p-adaptivité*), et d'une adaptation locale de la taille des éléments (*h-adaptivité*) commune aux EF (Etienne *et al.*, 2010a). Elle est de plus naturellement adaptée pour modéliser les phénomènes discontinus tels que la propagation d'ondes générées au cours de la rupture d'une faille (Ben Jemaa *et al.*, 2009; De la Puente *et al.*, 2009), ou encore pour modéliser les contacts fluide-solide.

Principes de base

Les éléments finis considèrent que la solution $\hat{\mathbf{u}}_i(x)$ peut être approximée en moyenne sur

l'élément i (ou cellule de contrôle) de position x , par une combinaison linéaire de fonctions de base (des polynômes de Lagrange φ_{ij}) définies au niveau de noeuds x_j (base nodale), de sorte que

$$\hat{\mathbf{u}}_i(x) = \sum_{j=1}^{d_i} \mathbf{u}_{ij}(x_j) \varphi_{ij}(x), \quad (1.17)$$

où d_i est le nombre de degrés de liberté de la cellule i , et les $\mathbf{u}_{ij}(x_j)$ représentent les degrés de libertés correspondant aux valeurs de \mathbf{u}_i aux noeuds x_j . La Figure 1.5(c-e) illustre que la représentation d'un sinus par des polynômes d'ordre P_k est plus précise en augmentant P_k , la solution étant approximée par une succession de valeurs constantes en P_0 , de droites en P_1 , et de paraboles en P_2 . On se rend compte sur la Figure 1.5a que la précision accrue procurée par un ordre élevé se fait au détriment d'un plus grand nombre de degrés de libertés. Pour un système d'ordre 1 où cinq champs sont considérés, le nombre de degrés de liberté pour une cellule triangulaire est de 5 en P_0 (1 noeud), 15 en P_1 (3 noeuds) et 30 en P_2 (6 noeuds). Précisons que la méthode nécessite un minimum de dix points par longueur d'onde à l'ordre P_0 et P_1 , et de 4 points à l'ordre P_2 .

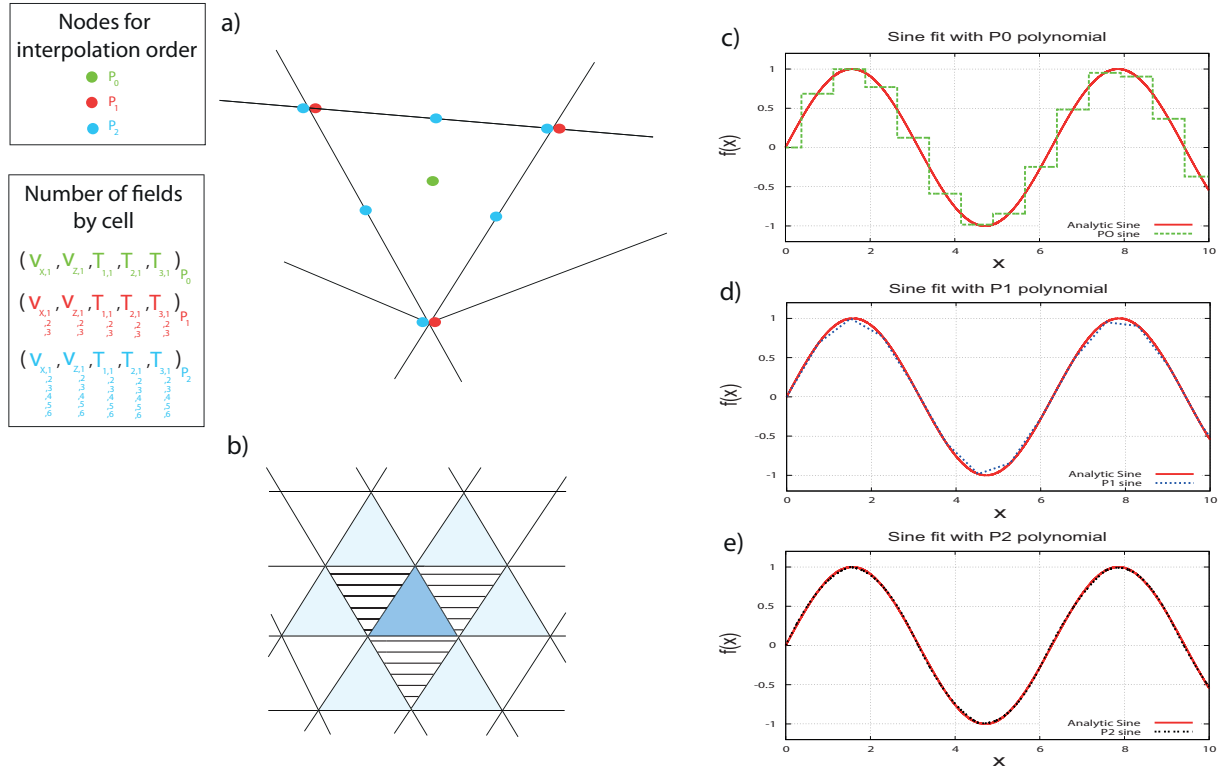


FIGURE 1.5 – (a) Cellule triangulaire sur laquelle sont localisés les noeuds associés aux interpolations P_0 (en vert), P_1 (en rouge) et P_2 (en bleu). Une solution en vitesses et contraintes doit être obtenue en chaque noeud. (b) Schéma numérique pour un système d'ordre 1 : 4 cellules dont 1 centrale et 3 cellules voisines, et pour un système d'ordre 2 (10 cellules en tout). (c-e) Approximation d'une sinusoïde par un polynôme d'ordre 0 (c), d'ordre 1 (d) et d'ordre 2 (e) (extrait de Brossier (2009)). Le nombre de degrés de liberté reste constant pour les trois figures.

L'approche des EF pour résoudre un système d'équation tel que celui de l'élastodynamique

(1.8 ou 1.5), consiste à remplacer les champs $\hat{\mathbf{u}}_i$ par leur expression donnée par l'équation 1.17, à multiplier de part et d'autre le système par une fonction *test* (φ_{ir}), identique aux fonctions de base (φ_{ij}), puis à intégrer spatialement sur la cellule le système pour moyenner la solution (approche *variationnelle*). Dans l'approche GD, la fonction test est nulle pour toutes les cellules autres que celle pour laquelle elle est définie, assurant ainsi le caractère discontinu de la méthode. Le système étant composé de dérivées spatiales du champ \mathbf{u} , une intégration par partie est réalisée du type :

$$\int \varphi_{ir}(x) \partial_x \hat{\mathbf{u}}_i(x) dx = \left[\varphi_{ir}(x) \hat{\mathbf{u}}_i(x) \right] - \int \partial_x \varphi_{ir}(x) \hat{\mathbf{u}}_i(x) dx . \quad (1.18)$$

Il peut être montré que le premier terme du membre de droite correspond à un flux numérique, et traduit le fait que les cellules, bien que disconnectées (noeuds non communs), communiquent entre elles via ces flux (flèches bleues sur la Figure 1.1b). L'intérêt d'avoir intégré le système puis fait une dérivation par partie réside dans l'élimination de la dérivée spatiale du champ, déplacée sur les fonctions de base. Une fois que les champs $\hat{\mathbf{u}}_i$ dans l'équation 1.18, sont remplacés par la somme sur les degrés de libertés \mathbf{u}_{ij} , ces derniers peuvent être sortis de l'intégrale. De plus, la dérivation de fonctions polynomiales est triviale, et l'intégration (de produits) de fonctions de base polynomiales est très simple et peu coûteuse en GD (petites matrices de taille $d_i \times d_i$ à calculer), grâce à des règles de quadrature de type Gauss-Legendre. L'approche des EF continus ne dispose pas de la *p-adaptativité* car les termes de flux sont uniquement définis aux limites du domaine (et non pas entre chaque cellule). Le caractère discontinu de la méthode GD permet de résoudre le problème localement au niveau des éléments de base. Par ailleurs, le code utilisé a été implémenté à partir d'un formalisme simplifié des GD dans lequel les propriétés physiques sont constantes au sein de l'élément (Brossier, 2011b). Cela peut cependant poser problème aux ordres élevés pour lesquels le nombre de cellules par longueur d'onde est plus faible. Précisons que (seul) l'ordre d'interpolation P_0 doit être appliqué sur des maillages réguliers triangulaires (de type équilatéral par exemple) pour assurer une précision acceptable des solutions (Brossier *et al.*, 2008b). L'ordre P_0 de la méthode GD est équivalent à une formulation en volumes finis à l'ordre P_0 , pour laquelle la fonction test est une constante (LeVeque, 2002), ce qui annule le deuxième terme de l'équation 1.18. Il a été montré que la formulation P_0 en volume fini appliqué sur une grille carrée est équivalente à la méthode des DF.

En GD, les termes de flux numériques font intervenir les voisins de la cellule *centrale* considérée (en bleu dans la Figure 1.5b). Lorsque le système est d'ordre 1 (en vitesses-contraintes), seuls les trois voisins directs de la cellule centrale sont impliqués (cellules hachurées), alors que le système en vitesses obtenu par une approche parsimonieuse fait aussi intervenir les voisins des cellules voisines (cellules bleu clair). Nous utilisons le code GD décrit dans Brossier (2011a), dans lequel est implémenté un système d'ordre 1 pour le calcul des champs, et d'ordre 2 pour la construction du gradient (voir partie 1.2.3.2). L'équation $\mathbf{A} \mathbf{p} = \mathbf{s}$ est alors formée pour le système d'ordre 1 d'une matrice d'impédance carrée d'ordre $K = 5$ (nombre de champs) \times $ddl \times N$, où ddl est le nombre de degrés de liberté par cellule (3 en P_1), et N le nombre de cellules. Dans le cas où l'ordre P_1 est utilisé pour toutes les cellules, la convention (arbitraire) utilisée est de mettre le vecteur \mathbf{P} sous la forme :

$(V_{x,j=1,3,1}, V_{z,j=1,3,1}, T_{1,j=1,3,1}, T_{2,j=1,3,1}, T_{3,j=1,3,1}, \dots, V_{x,j=1,3,N}, V_{z,j=1,3,N}, T_{1,j=1,3,N}, T_{2,j=1,3,N}, T_{3,j=1,3,N})^t$, où les indices $j = 1, 3$ dénotent une boucle sur les degrés de liberté, et où le troisième indice est le numéro de la cellule. Le nombre de coefficients non nuls pour les lignes associées à un noeud de contrainte est de :

$N_{T_i} = 1$ (terme de collocation associé à la contrainte) + 4 (1 cellule centrale + 3 cellules voi-

sines) $\times 2$ (pour V_X et V_Z) \times ddl = 25 en P_1 , alors que pour un noeud de vitesse le nombre de termes non nul vaut :

$N_{V_i} = 1 + 4 \times 3$ (pour T_1, T_2, T_3) \times ddl = 37 en P_1 . Le nombre total de coefficients non nuls dans la matrice \mathbf{A} est donc de :

[ddl $\times 3$ (pour T_1, T_2, T_3) $\times N_{T_i}$ + ddl $\times 2$ (pour V_X, V_Z) $\times N_{V_i}$] $\times N = 447 N_{P_1}$ en P_1 , et 53 N_{P_0} en P_0 (où N_{P_i} est le nombre de cellules d'ordre P_i). Pour donner un exemple, dans notre application la plus coûteuse en mémoire (Valhall élastique, voir Annexe B), nous avons utilisé près de 6 millions et 325000 cellules dans lesquelles est réalisé une interpolation P_0 et P_1 respectivement. La mémoire nécessaire pour stocker les coefficients non nuls associés est donc de : (53 N_{P_0} + 447 N_{P_1}) $\times 8$ octets = 3.7 Go.

La matrice d'impédance formée n'est pas bande-diagonale comme en différences finies en raison de la numérotation arbitraire des cellules i dans le vecteur \mathbf{p} . La phase d'analyse de MUMPS permet de réordonnancer les coefficients non nuls pour optimiser les opérations de pivots et le nombre de facteurs LU.

Des détails sur le temps de calcul et la mémoire nécessaire aux tests réalisés avec différentes approximations de l'équation d'onde, avec le code de différence finie et de Galerkin Discontinu, sont donnés dans l'Annexe B.

1.2 Problème inverse

1.2.1 Introduction

Un problème inverse consiste à estimer les valeurs de paramètres physiques (ou inconnues) à partir d'observations du phénomène physique. La démarche consiste à trouver la meilleure combinaison de paramètres (les solutions) qui permettent de rendre compte des mesures. Une fonction coût est alors définie pour traduire l'écart entre les données observées et les données calculées dans un modèle censé expliquer les données. On retrouve les problèmes inverses dans de nombreux domaines scientifiques tels que la géophysique interne, les prévisions météo, l'imagerie médicale, en traitement d'image, pour la détermination des constantes d'une réaction chimique, l'acoustique sous-marine ou encore en cosmologie. De plus, un nombre innombrable d'objets de la vie quotidienne a été conçu suite à la résolution d'un problème inverse. Un problème de thermique est par exemple résolu pour déterminer la forme des résistances d'un grille-pain de manière à ce que le pain soit grillé uniformément. L'acoustique architecturale permet d'optimiser la forme d'une salle de concert pour que le son soit perçu au mieux par toute l'assemblée. On peut encore citer la détermination de la forme d'un obstacle par un radar, ou bien les murs antibruit sur l'autoroute...

Un problème inverse est bien posé au sens d'Hadamard (1902) s'il a les propriétés suivantes :

- il existe une solution
- il y a unicité de la solution
- la solution dépend de façon continue des données du problème

Si l'une de ces trois conditions n'est pas satisfaite, alors le problème est mal posé, ce qui est quasiment toujours le cas. Si la première condition n'est pas satisfaite, un moyen d'y faire face est de considérer un problème plus simple dont la solution existe et dont l'introduction dans le problème sans solution permet de le résoudre de manière approchée (principe de relaxation). Le deuxième point est souvent lié à un nombre trop faible de données par rapport au nombre de paramètres (problème *sous-déterminé*), ou bien à cause d'une mauvaise distribution des données. Le moyen de choisir parmi les différentes solutions possibles est d'introduire de l'information a priori ou bien d'introduire des contraintes par exemple. Le troisième point, le plus problématique, est lié à l'insensibilité des paramètres aux données, par exemple lorsqu'une petite perturbation dans les données doit être expliquée par une importante perturbation de modèle. La non-linéarité d'un problème est accentuée par la présence de bruit dans les données, car il est difficile d'approcher de façon satisfaisante de la solution du problème inverse, ou bien lorsque la discrétisation numérique d'un problème est trop grossière pour rendre compte de phénomènes continus à plus petites échelles (troncature d'échelle en temps et/ou en espace). De la même manière, un modèle physique incorrect ou qui ne permet pas de rendre compte de la physique réelle des phénomènes avec suffisamment de précision, rend le problème inverse mal posé.

Afin de trouver la solution la plus convenable au problème considéré, il convient d'explorer l'espace des solutions. Pour cela, on distingue l'inversion par optimisation globale, qui fait une recherche systématique plus ou moins aléatoire de tous les modèles possibles (méthode de Monte Carlo par exemple) de l'inversion par optimisation locale. Cette dernière approche recherche de manière itérative un nouveau modèle au voisinage du modèle initial, qui minimisera la fonction coût. L'inversion des formes d'onde ne peut raisonnablement pas envisager une inversion globale

en raison de son coût numérique prohibitif, nous considérons par conséquent une inversion par optimisation locale. Celle-ci nécessite en contrepartie un modèle initial suffisamment proche du vrai modèle, afin de converger vers le minimum global de la fonction coût.

1.2.2 Théorie des méthodes locales d'optimisation

1.2.2.1 Equations normales des approches linéaires et non-linéaires

La relation entre les données \mathbf{d}_{obs} enregistrées lors d'une acquisition sismique et les paramètres du sous-sol est intrinsèquement non linéaire, et s'exprime via l'opérateur non linéaire \mathbf{g} comme $\mathbf{d}_{obs} = \mathbf{g}(\mathbf{m})$ (Tarantola, 1987), où \mathbf{m} est le vecteur modèle décrivant les paramètres considérés (fonction de la discrétisation du milieu et des classes de paramètres considérées). Nous allons maintenant distinguer deux approches pour appréhender ce problème : *l'inversion non-linéaire* consistant à linéariser le problème inverse, et *l'inversion linéaire* où le problème direct est linéarisé. L'inversion des formes d'ondes est basée sur la première approche pour imager des structures fines, alors que la seconde approche est par exemple utilisée pour des migrations de type ray-Born (Jin *et al.*, 1992) ou Kirchhoff (Keho, 1988; Thierry & Baina, 2001), pour lesquelles il y a une séparation d'échelle entre le macromodèle et la réflectivité. Nous allons dériver les équations normales associées à ces deux approches, puis nous comparerons le principe des processus itératifs non-linéaire et linéaire associés.

Linéarisation du problème inverse

Dans nos applications d'inversion des formes d'ondes, nous calculons le champ d'onde complet, exact au sens de la validité du problème direct choisi

$$\mathbf{d}_{cal} = \mathbf{g}(\mathbf{m}). \quad (1.19)$$

Définissons l'écart $\Delta\mathbf{d}$ entre les données enregistrées et les données calculées dans un modèle \mathbf{m}

$$\Delta\mathbf{d} = \mathbf{d}_{cal}(\mathbf{m}) - \mathbf{d}_{obs}. \quad (1.20)$$

Le modèle \mathbf{m} est construit par addition du modèle initial \mathbf{m}_0 de l'inversion, et de la perturbation du modèle $\delta\mathbf{m}$ calculée suite à l'inversion du problème. Ainsi, les données calculées peuvent s'écrire : $\mathbf{d}_{cal} = \mathbf{g}(\mathbf{m}) = \mathbf{g}(\mathbf{m}_0 + \delta\mathbf{m})$. Le critère utilisé pour mesurer la distance entre données calculées et observées est la fonction coût \mathcal{C} , définie par la norme Euclidienne L_2 de l'écart $\Delta\mathbf{d}$

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \Delta\mathbf{d}^\dagger \Delta\mathbf{d}, \quad (1.21)$$

où \dagger indique l'opérateur adjoint (transposée complexe). La linéarisation du problème inverse consiste à faire un développement de Taylor au second ordre de la fonction coût au voisinage de \mathbf{m}_0 , en supposant que les perturbations de modèle $\delta\mathbf{m}$ sont suffisamment petites par rapport au vecteur modèle \mathbf{m}

$$\mathcal{C}(\mathbf{m}) = \mathcal{C}(\mathbf{m}_0) + \frac{\partial\mathcal{C}(\mathbf{m}_0)}{\partial\mathbf{m}}\delta\mathbf{m} + \frac{1}{2} \frac{\partial^2\mathcal{C}(\mathbf{m}_0)}{\partial\mathbf{m}^2}\delta\mathbf{m}^2 + \mathcal{O}(\mathbf{m}^3). \quad (1.22)$$

Dans un problème inverse non-linéaire, la fonction coût $\mathcal{C}(\mathbf{m})$ n'est pas supposée être quadratique (polynôme d'ordre 2), mais '*localement*' quadratique. Le problème inverse consiste à

réaliser un processus itératif en linéarisant la fonction coût à chaque itération, comme illustré par les deux paraboles en noir pointillé sur la Figure 1.6, qui tangent la fonction $\mathcal{C}(\mathbf{m})$ au niveau des points \mathbf{m}_0 et \mathbf{m}_1 . Nous différencions à présent la relation 1.22 par rapport à \mathbf{m} en négligeant les termes d'ordres 3 qui sont nuls localement autour de \mathbf{m}_0

$$\frac{\partial \mathcal{C}(\mathbf{m})}{\partial \mathbf{m}} = \frac{\partial \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}} + \frac{\partial^2 \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}^2} \delta \mathbf{m}.$$

Le minimum de la fonction coût *locale* (parabole noire pointillée sur la Figure 1.6) est atteint à proximité du point \mathbf{m}_0 lorsque la dérivée première de la fonction coût s'annule $\partial \mathcal{C}(\mathbf{m})/\partial \mathbf{m} = 0$. La perturbation de modèle $\delta \mathbf{m}_1$ à l'itération 1 est alors obtenue par l'expression

$$\delta \mathbf{m}_1 = - \left[\frac{\partial^2 \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}^2} \right]^{-1} \frac{\partial \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}}. \quad (1.23)$$

Le terme $\partial \mathcal{C}(\mathbf{m}_0)/\partial \mathbf{m}$ est le gradient de la fonction coût, définissant sa direction de plus forte croissance au point \mathbf{m}_0 . La dérivée seconde de la fonction coût est appelée le Hessian, et décrit l'inflexion de la fonction à proximité du point \mathbf{m}_0 . Par ailleurs, on se rend compte sur la Figure 1.6 que l'inversion peut facilement être piégée dans un minimum local si elle part d'un modèle initial trop éloigné du minimum global. Dans le cadre d'un processus itératif non linéaire, le modèle à l'itération courante ' k ' est relié au modèle de l'itération précédente par $\mathbf{m}^k = \mathbf{m}^{k-1} + \delta \mathbf{m}^k$, avec $\delta \mathbf{m}^k$ défini comme

$$\delta \mathbf{m}_k = - \left[\frac{\partial^2 \mathcal{C}(\mathbf{m}_{k-1})}{\partial \mathbf{m}^2} \right]^{-1} \frac{\partial \mathcal{C}(\mathbf{m}_{k-1})}{\partial \mathbf{m}}. \quad (1.24)$$

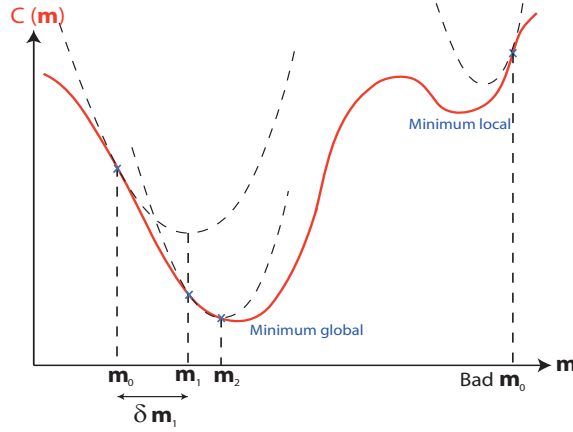


FIGURE 1.6 – Illustration d'un problème inverse non-linéaire itératif, procédant par linéarisations locales successives de la fonction coût au niveau des modèles \mathbf{m}_k .

Les dérivées premières et secondes de la fonction coût autour du modèle \mathbf{m}_0 peuvent s'exprimer, à l'aide des relations 1.19 et 1.20, comme (Menke, 1984; Tarantola, 1987; Pratt *et al.*, 1998)

$$\frac{\partial \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}} = -\Re \left[\left(\frac{\partial \mathbf{g}(\mathbf{m}_0)}{\partial \mathbf{m}} \right)^\dagger (\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_0)) \right] = -\Re [\mathbf{J}_0^t \Delta \mathbf{d}^*], \quad (1.25)$$

$$\frac{\partial^2 \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}^2} = \Re [\mathbf{J}_0^\dagger \mathbf{J}_0] + \Re \left[\left(\frac{\partial \mathbf{J}_0}{\partial \mathbf{m}} \right)^t (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right], \quad (1.26)$$

où l'opérateur \mathbf{J}_0 est la matrice des dérivées de Fréchet au point \mathbf{m}_0 , et définit la dérivée partielle des données par rapport à \mathbf{m}

$$\mathbf{J}_0 = \frac{\partial \mathbf{g}(\mathbf{m}_0)}{\partial \mathbf{m}}. \quad (1.27)$$

Le premier terme du Hessien $\Re (\mathbf{J}_0^\dagger \mathbf{J}_0)$ est appelé le Hessien approximé \mathbf{H}_a (Pratt *et al.*, 1998). En substituant ces expressions dans l'équation 1.23, nous trouvons les équations normales à l'itération k

$$\delta \mathbf{m}_k = \left\{ \Re [\mathbf{J}_{k-1}^\dagger \mathbf{J}_{k-1}] + \Re \left[\left(\frac{\partial \mathbf{J}_{k-1}}{\partial \mathbf{m}} \right)^t (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right] \right\}^{-1} \Re [\mathbf{J}_{k-1}^t \Delta \mathbf{d}_{k-1}^*]. \quad (1.28)$$

Linéarisation du problème direct

Afin de mieux situer l'inversion des formes d'ondes complètes parmi les autres méthodes d'imagerie, nous prenons pour exemple l'inversion / migration rai-Born pour illustrer les différences existant entre la technique de linéarisation du problème inverse et celle consistant à linéariser le problème direct. Dans ce cadre, la linéarisation du problème direct repose sur le fait qu'un modèle \mathbf{m} peut-être vu comme la somme d'un modèle de référence lisse (\mathbf{m}_0), et d'une perturbation de modèle $\delta \mathbf{m}$ associée aux hauts nombres d'ondes spatiaux (ou bien à la réflectivité) du modèle. Contrairement à l'inversion non-linéaire où l'on résout le champ d'onde complet au détriment d'un important coût numérique, l'inversion / migration rai-Born calcule une approximation de l'équation d'onde dans un milieu lisse via la théorie des rais, et relie linéairement les hauts nombres d'ondes spatiaux du modèle au champ diffracté par le modèle de perturbation (approximation de Born). La perturbation des données $\delta \mathbf{d}$ associée à la perturbation de modèle $\delta \mathbf{m}$ peut s'écrire

$$\delta \mathbf{d} = \mathbf{d} - \mathbf{d}_0 = g(\mathbf{m}_0 + \delta \mathbf{m}) - g(\mathbf{m}_0). \quad (1.29)$$

Par ailleurs, le développement de Taylor à l'ordre 1 de l'opérateur $g(\mathbf{m})$ prend la forme

$$g(\mathbf{m}) = g(\mathbf{m}_0 + \delta \mathbf{m}) = g(\mathbf{m}_0) + \frac{\partial g(\mathbf{m}_0)}{\partial \mathbf{m}} \delta \mathbf{m} + \mathcal{O}(\mathbf{m}^2). \quad (1.30)$$

Ignorer les termes de second ordre ($\mathcal{O}(\mathbf{m}^2)$) revient à négliger les diffractions doubles. On obtient alors

$$g(\mathbf{m}_0 + \delta \mathbf{m}) - g(\mathbf{m}_0) \approx \frac{\partial g(\mathbf{m}_0)}{\partial \mathbf{m}} \delta \mathbf{m}, \quad (1.31)$$

ce qui en utilisant les équations 1.27 et 1.29 aboutit à la relation de Born, reliant linéairement la perturbation des données à la perturbation du modèle

$$\delta \mathbf{d} = \mathbf{J}_0 \delta \mathbf{m}. \quad (1.32)$$

On peut maintenant exprimer l'écart entre les données observées et calculées comme

$$\Delta \mathbf{d} = \mathbf{d}_{obs} - g(\mathbf{m}) = \mathbf{d}_{obs} - g(\mathbf{m}_0) - \mathbf{J}_0 \delta \mathbf{m} = \delta \mathbf{d}_{obs} - \delta \mathbf{d}_{cal}. \quad (1.33)$$

Le terme $\delta \mathbf{d}_{obs} = \mathbf{d}_{obs} - g(\mathbf{m}_0)$ représente les données observées, mutées de sorte à enlever les ondes directes et les arrivées grands angles, et prétraitées de sorte à enlever les données issues de diffractions et réflexions multiples. Le terme $\delta \mathbf{d}_{cal} = \mathbf{J}_0 \delta \mathbf{m}$ est associé aux diffractions manquantes dans le modèle m_0 , c'est-à-dire aux arrivées réfléchies dont nous voulons extraire la réflectivité du milieu. Les dérivées de Fréchet \mathbf{J}_0 sont calculées dans le milieu lisse \mathbf{m}_0 via la théorie des rais. Dans l'approximation de Born, $\delta \mathbf{d}_{cal}$ est formé de la contribution de l'ensemble des champs diffractés par une grille de points diffractants, lesquels sont indépendants les uns des autres (pas de diffractions multiples). Alors que dans l'approximation de Kirchhoff, les diffractants sont reliés de manière continue et forment une interface qui réfléchit globalement les ondes avec une intensité fonction du coefficient de réflexion.

Nous avons considéré que la fonction $g(\mathbf{m})$ est d'ordre 1 (équation 1.30), ce qui implique que la dérivée seconde de g par rapport à \mathbf{m} est nulle. Ainsi dans cette approche, la dérivée première de la fonction coût prendra la même forme que pour l'inversion non-linéaire (équation 1.25), mais la dérivée seconde de la fonction coût ne comporte plus que le premier terme de l'équation 1.26, de sorte que

$$\frac{\partial^2 \mathcal{C}(\mathbf{m}_0)}{\partial \mathbf{m}^2} = \Re \left[\mathbf{J}_0^\dagger \mathbf{J}_0 \right]. \quad (1.34)$$

La différence entre les deux expressions du Hessien est le terme $\Re \left[\frac{\partial \mathbf{J}_0^t}{\partial \mathbf{m}^t} (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right]$. Ce terme est en effet associé aux diffractions doubles (voir partie 1.2.3.3) (Pratt *et al.*, 1998), non prises en compte dans les problèmes linéaires car les termes de second ordre de $g(\mathbf{m})$ sont négligés.

Itérations linéaires versus non-linéaires

L'approximation de Born repose sur l'hypothèse que le problème est linéaire (équation 1.32), le problème est donc supposé converger en une itération. Cependant, un processus itératif est en général réalisé afin de compenser les effets de bande passante limitée du signal, de l'illumination non complète du milieu et l'approximation de diffraction simple du problème direct. Pour cela, un processus itératif *linéaire* est appliqué, dans lequel le modèle initial est constant à toutes les itérations et où les dérivées de Fréchet ne sont calculées qu'une seule fois dans le modèle \mathbf{m}_0 . En utilisant les équations normales d'un problème linéaire, la perturbation de modèle à l'itération k s'exprime comme

$$\begin{aligned} \mathbf{m}_k &= \mathbf{m}_0 + \Re \left[\mathbf{J}_0^\dagger \mathbf{J}_0 \right]^{-1} \Re \left[\mathbf{J}_0^\dagger (\delta \mathbf{d}_{obs} - \delta \mathbf{d}_{cal,k-1}) \right] \text{ avec} \\ \delta \mathbf{d}_{cal,k-1} &= \mathbf{J}_0 \sum_{i=1}^{k-1} \delta \mathbf{m}_i, \end{aligned}$$

où N_{it} est le nombre total d'itérations. Les conditions initiales de ce processus itératif sont $\delta \mathbf{m}_0 = 0$ et $\delta \mathbf{d}_{cal,0} = 0$, ce qui donne un gradient à la première itération valant $\Re \left[\mathbf{J}_0^\dagger \delta \mathbf{d}_{obs} \right]$. Nous pouvons faire le parallèle entre cette expression et l'idée développée par Lailly (1983a) et Tarantola (1984a), proposant de voir la migration comme la première itération d'un problème inverse. La migration est en effet basé sur la corrélation d'un champ incident avec le champ des données observées ($\delta \mathbf{d}_{obs}$) rétropropagé des récepteurs au point diffractant (voir section 1.2.3.2).

Par comparaison, nous rappelons le principe d'un processus itératif non-linéaire, associés

aux problèmes non-linéaires

$$\begin{aligned}\mathbf{m}_k &= \mathbf{m}_{k-1} + \Re \left[\mathbf{J}_{k-1}^\dagger \mathbf{J}_{k-1} + \left(\frac{\partial \mathbf{J}_{k-1}}{\partial \mathbf{m}} \right)^t (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right]^{-1} \Re \left[\mathbf{J}_{k-1}^\dagger (\mathbf{d}_{obs} - \mathbf{d}_{cal,k-1}) \right] \text{ avec} \\ \mathbf{d}_{cal,k-1} &= g(\mathbf{m}_{k-1}).\end{aligned}$$

1.2.2.2 Algorithmes d'optimisation locale

Nous présentons à présent les différents algorithmes d'optimisation qui conduisent plus ou moins rapidement le processus itératif vers le minimum global de la fonction coût, pour peu que le modèle initial soit localisé dans le bassin d'attraction de celui-ci. Tout d'abord, la méthode de **Newton** consiste à calculer le Hessien complet, soit les deux termes de l'équation 1.26, alors que la méthode de **Gauss-Newton** calcule de manière complète le premier terme du Hessien (Pratt *et al.*, 1998; Métivier *et al.*, 2012a). Pour les problèmes de grandes tailles, il est trop coûteux d'estimer l'un ou l'autre de ces termes, bien que des études utilisant ces méthodes aient déjà été présentées (Akçelik, 2002; Askan *et al.*, 2007). Nous allons maintenant décrire les approches qui diffèrent dans la manière dont est approximé le Hessien.

Méthode du gradient

Dans la méthode du gradient, le modèle à l'itération k est obtenu en suivant la ligne de plus forte décroissance de la fonction coût, via l'expression

$$\mathbf{m}_k = \mathbf{m}_{k-1} - \alpha_k \mathcal{G}_k, \quad (1.35)$$

où \mathcal{G} est le gradient de la fonction coût, et α est un scalaire réel appelé pas de descente, définissant l'amplitude de la perturbation à appliquer au modèle. La non prise en compte du Hessien implique que la direction de descente ne prend pas en compte la courbure de la fonction coût, et que le gradient ne sera pas corrigé des effets d'atténuation géométrique (voir partie 1.2.3.3). Cette approche convergera donc moins vite qu'une méthode prenant en compte une approximation plus fine du Hessien, et a plus de chances de converger vers un minimum local.

Pas de descente

Une des manières d'estimer le pas de descente consiste à faire une approximation linéaire (Tarantola, 1987) de sorte la fonction coût soit minimisée le long de la direction du gradient. Pour les problèmes multiparamètres, une méthode des sous-espaces a été développée par (Sambridge *et al.*, 1991) pour estimer un pas de descente pour chaque classe de paramètre. Cependant, la validité de l'approche peut-être affectée par la non-linéarité du problème et la non-séparabilité des paramètres. Une autre stratégie, utilisée dans mes applications, est d'estimer par ajustement parabolique le pas de descente. Pour cela, trois valeurs de fonction coût associées à un pas de descente sont évaluées. Une première valeur est fournie par le modèle initial, correspondant à $\alpha_0 = 0$. Les deux autres fonctions coût sont calculées pour des pas de descente supérieurs, et doivent satisfaire les conditions : $C(\alpha_1) < C(0)$ et $C(\alpha_1) < C(\alpha_2)$. Les trois points obtenus permettent de faire une interpolation parabolique dans le plan ($C - \alpha$) (Figure 1.7), donnant le pas de descente optimum (α_{opt}) minimisant la fonction coût Tarantola (1987).

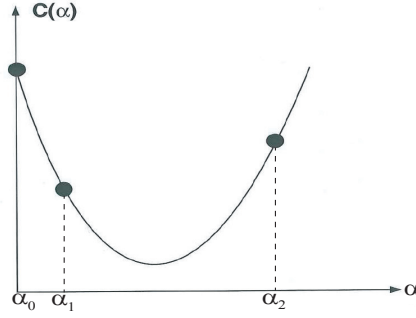


FIGURE 1.7 – Illustration de l'estimation du pas de descente α par ajustement parabolique. $C(\alpha)$ représente la fonction coût.

Gradient conjugué

La convergence de la méthode du gradient peut être accélérée grâce à la méthode du gradient conjugué, laquelle définit une direction de descente *conjuguée* qui optimise la recherche dans l'espace des modèles en combinant les directions anciennes et courante (Mora, 1987, 1988). Le processus itératif est alors supposé converger en N_M itérations, avec N_M le nombre de paramètres du vecteur modèle. D'après (Polak & Ribière, 1969), la direction de descente conjuguée \mathbf{p}_k à l'itération k est donnée par

$$\mathbf{p}_k = \mathcal{G}_{\mathbf{m},k-1} + \beta_k \mathbf{p}_{k-2}, \quad \text{avec} \quad \beta_k = \frac{(\mathcal{G}_{\mathbf{m},k-1} - \mathcal{G}_{\mathbf{m},k-2})^t \mathcal{G}_{\mathbf{m},k-1}}{\mathcal{G}_{\mathbf{m},k-2}^t \mathcal{G}_{\mathbf{m},k-2}}. \quad (1.36)$$

L-BFGS

L'algorithme d'optimisation *BFGS*, pour *Broyden-Fletcher-Goldfarb-Shanno* du nom de ses auteurs, est une méthode de type quasi-Newton qui permet de faire une *estimation* économique du Hessien ou de son inverse \mathcal{H} (grâce à la formule de Sherman-Morrison-Woodbury), à partir des gradients et des modèles des itérations précédentes (Nocedal & Wright, 1999). Le principe de l'approche consiste à utiliser les variations du gradient d'une itération à l'autre pour construire par différences finies un modèle de la fonction coût. La version L-BFGS pour *Limited-BFGS*, n'utilise qu'un nombre limité d'itérations précédentes (5 dans nos applications) (Nocedal, 1980). La perturbation de modèle à l'itération k est directement estimée par $\delta \mathbf{m}_k = -\mathcal{H}_k \mathcal{G}_k$ à partir d'additions, de différences et de produits scalaires de vecteurs sans avoir à stocker la matrice Hessienne. Il est possible de fournir à l'algorithme du L-BFGS une estimation initiale du Hessien, en normalisant le gradient $\tilde{\mathcal{G}}$ par la diagonale du Hessien approximé (\mathbf{H}_a) (Pratt *et al.*, 1998; Operto *et al.*, 2006)

$$\tilde{\mathcal{G}}_k = \frac{\mathcal{G}_k}{\Re[\text{diag}(\mathbf{H}_a)] + \epsilon}, \quad (1.37)$$

pour une convergence plus rapide et meilleure. ϵ est un terme d'amortissement de l'inversion (voir partie 1.2.3.3). Les modèles fournis au L-BFGS sont normalisés par une valeur de référence associée à chaque classe de paramètre, car les produits scalaires (associés à une norme) réalisés dans l'algorithme requièrent des vecteurs adimensionnels (Brossier, 2011b). Les perturbations de modèles obtenues à l'issue de l'algorithme sont ensuite dénormalisées pour rendre son unité au paramètre. L'estimation du Hessien complet par le L-BFGS inclut les termes hors-diagonaux

du Hessian complet et rend compte des effets de diffractions doubles, sans avoir à payer le prix d'une véritable itération de Newton ou Gauss-Newton. La capacité du L-BFGS à prendre en compte les diffractions doubles est illustrée sur la Figure 1.8, extraite de Métivier *et al.* (2012a). La reconstruction par inversion des formes d'ondes avec le L-BFGS est bien meilleure qu'avec un algorithme de Gauss-Newton pour distinguer deux hétérogénéités distantes de moins d'une demi-longueur d'onde, et se trouve être un peu moins précise qu'avec la méthode de Newton (estimation exacte du Hessian complet). Le L-BFGS s'avère aussi être crucial dans le cadre d'inversion multiparamètres pour mettre à l'échelle les différentes classes de paramètres (Brossier, 2011b). Nous reviendrons plus en détail sur les multiples rôles du Hessian dans la partie 1.2.3.3. Dans la pratique, nous couplons l'utilisation du L-BFGS avec *un seul* pas de descente pour toutes les classes de paramètre (voir l'algorithme de FWI sur la Figure 1.12), la pondération de la perturbation associée à chaque classe de paramètre étant déjà estimée par le L-BFGS. Brossier *et al.* (2009b) a montré que l'application du L-BFGS en inversion des formes d'ondes sur un cas synthétique s'est révélé supérieure au gradient conjugué en terme de rapidité de convergence et de minimum atteint par la fonction coût.

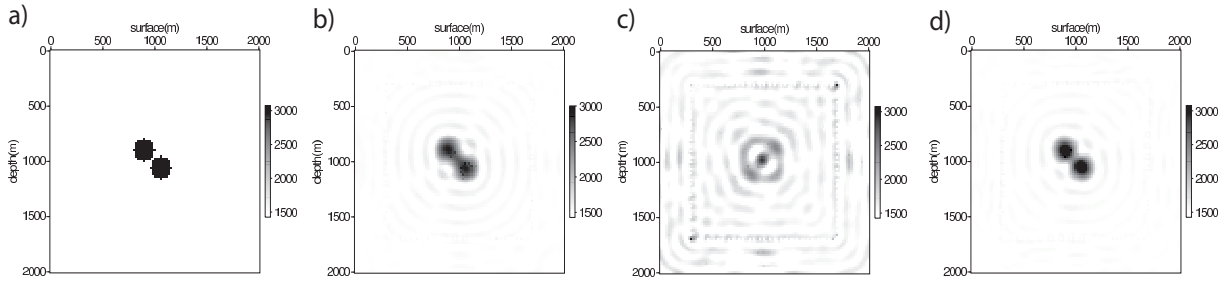


FIGURE 1.8 – Inversion des formes d'onde acoustique du paramètre V_P . (a) Modèle synthétique *vrai*, comportant deux inclusions circulaires de 190 m de diamètre. Les bords des inclusions sont espacés d'un minimum de 40 m, sachant que la longueur d'onde est de 300 m (à 5 Hz). (b-d) Résultats obtenu avec le L-BFGS (b), avec un algorithme de type Gauss-Newton (c) et de Newton (d). Les termes de diffractions doubles sont pris en compte avec le L-BFGS, contrairement à l'algorithme de type Gauss-Newton. Extrait de Métivier *et al.* (2012a).

1.2.2.3 Calcul du gradient d'une fonctionnelle par la méthode de l'état adjoint

Cette partie reprend les notations et le formalisme développé par Plessix (2006), dans lequel de plus amples détails pourront être trouvés.

Nous désignons par \mathbf{U} l'espace de la variable d'état u , \mathbf{U}^\dagger l'espace adjoint de la variable d'état adjointe λ (ou multiplicateur de Lagrange), \mathbf{M} l'espace des modèles m et \mathbf{R} l'espace des réels. La variable d'état doit satisfaire l'équation d'état (ou problème direct)

$$F(u(m), m) = 0, \quad (1.38)$$

et nous définissons la fonctionnelle \mathcal{C} via la fonctionnelle h comme

$$\mathcal{C}(m) = h(u(m), m). \quad (1.39)$$

Par ailleurs, nous allons montrer que l'étude de la fonctionnelle augmentée \mathcal{L} , appelée le Lagrangien, va nous permettre de dériver l'expression du gradient de la fonctionnelle \mathcal{C} que nous

voulons minimiser (ou maximiser), sous la contrainte de satisfaire l'équation d'état 1.38. \mathcal{L} est défini comme

$$\mathcal{L}(\tilde{u}, \tilde{\lambda}, m) = h(\tilde{u}, m) - \left\langle \tilde{\lambda}, F(\tilde{u}, m) \right\rangle_{\mathbf{U}}, \quad (1.40)$$

où $\langle \cdot \rangle_{\mathbf{U}}$ est le produit scalaire défini dans l'espace \mathbf{U} . Une variable est écrite \tilde{u} (ou $\tilde{\lambda}$) si elle peut prendre n'importe quelle valeur appartenant à \mathbf{U} , et u (ou λ) si elle doit satisfaire à une réalisation physique telle que l'équation d'état 1.38. Nous notons l'opération de passage de l'espace de départ \mathbf{A} vers l'espace d'arrivée \mathbf{B} via l'opérateur C comme $C : \mathbf{A} \rightarrow \mathbf{B}$. Les opérateurs précédemment mentionnés ont pour espace de départ et d'arrivée

$$F : \mathbf{U} \times \mathbf{M} \rightarrow \mathbf{U} ; \quad h : \mathbf{U} \times \mathbf{M} \rightarrow \mathbf{R} ; \quad \mathcal{C} : \mathbf{M} \rightarrow \mathbf{R} ; \quad \mathcal{L} : \mathbf{U} \times \mathbf{U}^\dagger \times \mathbf{M} \rightarrow \mathbf{R}. \quad (1.41)$$

Nous allons maintenant dériver \mathcal{L} en fonction de ses variables \tilde{u} , $\tilde{\lambda}$ et m , en considérant que \tilde{u} et $\tilde{\lambda}$ sont indépendants de m , et nous allons regarder la forme que prennent ces équations au niveau des variables u et λ qui doivent satisfaire les équations d'état. Tout d'abord nous dérivons le Lagrangien par rapport à λ

$$\frac{\partial \mathcal{L}(\tilde{u}, \tilde{\lambda}, m)}{\partial \tilde{\lambda}} = -F(\tilde{u}, m). \quad (1.42)$$

Nous choisissons u de sorte que cette dérivée soit nulle, c'est-à-dire lorsque l'équation d'état 1.38 est satisfaite. Au point u , on remarque que le Lagrangien s'écrit $\mathcal{L}(u, \tilde{\lambda}, m) = h(u, m) = \mathcal{C}(m)$. La différentiation en chaîne de cette équation, sachant que seuls u et (logiquement) m sont fonctions de m , donne

$$\frac{\partial \mathcal{L}(u, \tilde{\lambda}, m)}{\partial \tilde{u}} \frac{\partial u}{\partial m} + \frac{\partial \mathcal{L}(u, \tilde{\lambda}, m)}{\partial m} \frac{\partial m}{\partial m} = \frac{\partial \mathcal{C}}{\partial m}. \quad (1.43)$$

Par ailleurs, simultanément à la condition sur u (équation 1.38), nous choisissons un λ dans \mathbf{U}^\dagger qui satisfasse l'équation (en dérivant l'équation 1.40)

$$\frac{\partial \mathcal{L}(u, \lambda, m)}{\partial \tilde{u}} = 0 = \frac{\partial h(u, m)}{\partial \tilde{u}} - \left(\frac{\partial F(u, m)}{\partial \tilde{u}} \right)^\dagger \lambda. \quad (1.44)$$

Cette équation est l'équation d'état adjointe associée à la variable d'état adjointe λ . Lorsque les deux équations d'état sont satisfaites simultanément, c'est-à-dire au point selle du Lagrangien (où les deux dérivées de \mathcal{L} par rapport à u et λ s'annulent), l'équation 1.43 donne l'expression du gradient de la fonctionnelle \mathcal{C} (à l'aide de l'équation 1.40)

$$\frac{\partial \mathcal{C}}{\partial m} = \frac{\partial \mathcal{L}(u, \lambda, m)}{\partial m} \quad (1.45)$$

$$= \frac{\partial h(u, m)}{\partial m} - \left\langle \lambda, \frac{\partial F(u, m)}{\partial m} \right\rangle_{\mathbf{U}}. \quad (1.46)$$

Cette méthodologie basée sur les dérivées du Lagrangien permet ainsi d'obtenir la formule du gradient de n'importe quelle fonctionnelle devant satisfaire à un nombre quelconque de contraintes. Nous appliquerons dans la partie 1.2.3.2 cette approche pour obtenir le gradient de la fonction coût définie en FWI, et nous montrerons l'intérêt qu'elle présente.

1.2.3 Application à l'inversion des formes d'ondes

1.2.3.1 Approche fréquentielle

Le problème direct et le problème inverse sont effectués dans le domaine fréquentiel, bien que les deux problèmes puissent être découplés, en réalisant le problème direct en temps et le problème inverse en fréquence par exemple (*US PATENT NO.* 7,725, 266, et Nihei & Li (2007); Sirgue *et al.* (2008); Etienne *et al.* (2010b)). Nous effectuons l'inversion des données dans le domaine fréquentiel pour plusieurs raisons. L'inversion de quelques fréquences suffit à obtenir une image focalisée en raison de la couverture redondante des nombres d'onde dans l'espace des modèles. Celle-ci résulte du double contrôle de la fréquence f et de l'ouverture angulaire θ sur les nombres d'onde k du modèle (voir partie 1.2.3.4). De plus, l'inversion dans le domaine des fréquences s'effectue d'une manière hiérarchique en procédant successivement de l'inversion des basses fréquences vers les hautes fréquences, conférant à l'inversion un caractère multi-résolution précieux pour contrôler la non-linéarité du problème inverse (Pratt, 1990; Pratt & Worthington, 1990; Brossier *et al.*, 2009a; Virieux & Operto, 2009). En effet, l'inversion des basses fréquences, reconstituant les bas nombres d'ondes du modèle, est moins sensible aux imprécisions du modèle de départ que ne le sont les hautes fréquences (problème de saut de phase dit '*cycle skipping*'). L'inversion hiérarchique des basses fréquences vers les hautes fréquences permet de ce fait d'améliorer progressivement la résolution des modèles de départ au fur et à mesure que l'inversion progresse vers les hautes fréquences. A noter qu'une approche similaire fut proposée par Bunks *et al.* (1995) en domaine temporel. Afin d'illustrer la boucle sur les fréquences à l'origine du caractère multi-résolution de l'imagerie, nous décrirons dans la partie 1.2.3.5 la structure de base du code (Figure 1.12). Par ailleurs, ces quelques composantes fréquentielles peuvent être modélisées très efficacement dans le domaine fréquentiel pour un grand nombre de sources et un nombre limité de fréquences, grâce à l'utilisation d'un solveur direct de type factorisation LU. L'approche fréquentielle permet de plus une implémentation pratique des conditions aux limites du type PML ('Perfectly Matched Layers') (Berenger, 1994), ainsi que de l'atténuation pour la description des milieux visco-élastiques. Celle-ci est alors introduite via des vitesses complexes (voir chapitre 5), alors que des fonctions de relaxations, plus complexes à implémenter, sont introduites pour les modélisations en temps (Tarantola, 1988).

1.2.3.2 Construction du gradient par la méthode de l'état adjoint, et interprétation du gradient

Nous reprenons dans cette partie les notations introduites dans les parties 1.2.2.1 et 1.2.2.3. Dans la méthode d'inversion des formes d'ondes, nous considérons l'équation d'état (d'après l'équation 1.13)

$$F(\mathbf{p}(\mathbf{m}), \mathbf{m}) = \mathbf{A}(\mathbf{m})\mathbf{p}(\mathbf{m}) - \mathbf{s} = 0. \quad (1.47)$$

La fonctionnelle $\mathcal{C}(\mathbf{m})$ traduit l'écart entre les données observées \mathbf{d}_{obs} et les données calculées en résolvant l'équation 1.47. Elle ne contient ici par simplicité qu'un terme dans l'espace des données (contrairement à la régularisation Tikhonov, partie 1.2.3.6), et satisfait

$$\mathcal{C}(\mathbf{m}) = h(\mathbf{p}(\mathbf{m}), \mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^\dagger \Delta \mathbf{d} = \frac{1}{2} \|\mathbf{p}(\mathbf{m}) - \mathbf{d}_{obs}\|^2. \quad (1.48)$$

Nous considérons que $\tilde{\mathbf{p}}$ et $\tilde{\boldsymbol{\lambda}}$ sont indépendants de \mathbf{m} , contrairement à \mathbf{p} et $\boldsymbol{\lambda}$ qui doivent satisfaire les équations d'état pour un \mathbf{m} donné. Le Lagrangien associé à la fonctionnelle h et à la contrainte F s'écrit donc

$$\mathcal{L}(\tilde{\mathbf{p}}, \tilde{\boldsymbol{\lambda}}, \mathbf{m}) = h(\tilde{\mathbf{p}}) - \left\langle \tilde{\boldsymbol{\lambda}}, F(\tilde{\mathbf{p}}, \mathbf{m}) \right\rangle_{\mathbf{x}} \quad (1.49)$$

$$= \Re \left[\frac{1}{2} \|\tilde{\mathbf{p}} - \mathbf{d}_{obs}\|^2 - \left\langle \tilde{\boldsymbol{\lambda}}, \mathbf{A}(\mathbf{m})\tilde{\mathbf{p}} - \mathbf{s} \right\rangle_{\mathbf{x}} \right], \quad (1.50)$$

où $\langle \rangle_{\mathbf{x}}$ est un produit scalaire dans l'espace des complexes \mathbf{C}^n , avec n le nombre de points discrétisant le modèle. La dérivée de \mathcal{L} par rapport à $\tilde{\boldsymbol{\lambda}}$ redonne l'équation 1.47 du problème direct. Le champ adjoint $\boldsymbol{\lambda}$ annule la dérivée de \mathcal{L} par rapport à $\tilde{\mathbf{p}}$

$$\frac{\partial \mathcal{L}(\mathbf{p}, \boldsymbol{\lambda}, \mathbf{m})}{\partial \tilde{\mathbf{p}}} = 0 = \frac{\partial h(\mathbf{p}(\mathbf{m}))}{\partial \tilde{\mathbf{p}}} - \left(\frac{\partial F(\mathbf{p}(\mathbf{m}), \mathbf{m})}{\partial \tilde{\mathbf{p}}} \right)^{\dagger} \boldsymbol{\lambda} = \mathbf{R}_d^t \Delta \mathbf{d} - \mathbf{A}^{\dagger} \boldsymbol{\lambda}, \quad (1.51)$$

où \mathbf{R}_d^t projette le vecteur résidu $\Delta \mathbf{d}$ dans l'espace du problème direct. L'équation d'état adjoint s'écrit donc, en prenant le conjugué $*$ de l'équation 1.51, comme

$$\mathbf{A}^t \boldsymbol{\lambda}^* = \mathbf{R}_d^t \Delta \mathbf{d}^* \quad (1.52)$$

Une analogie avec l'équation 1.47 montre que cette équation correspond à la rétropropagation dans le milieu d'un champ émis par une source formée par l'ensemble des résidus. $\Delta \mathbf{d}$ est ici un vecteur associé à une seule source.

La dérivée de \mathcal{L} par rapport à \mathbf{m} donne le gradient de la fonctionnelle \mathcal{C} , en rappelant que $\tilde{\mathbf{p}}$ et $\tilde{\boldsymbol{\lambda}}$ sont indépendants de \mathbf{m}

$$\frac{\partial \mathcal{C}}{\partial \mathbf{m}} = \frac{\partial \mathcal{L}(\mathbf{p}, \boldsymbol{\lambda}, \mathbf{m})}{\partial \mathbf{m}} = - \left\langle \boldsymbol{\lambda}, \frac{\partial F(\mathbf{p}, \mathbf{m})}{\partial \mathbf{m}} \right\rangle_{\mathbf{x}} = -\Re \left[\left\langle \boldsymbol{\lambda}, \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right) \mathbf{p} \right\rangle_{\mathbf{x}} \right] \quad (1.53)$$

$$= -\Re \left[\left\langle (\mathbf{A}^{-1})^{\dagger} \mathbf{R}_d^t \Delta \mathbf{d}, \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right) \mathbf{p} \right\rangle_{\mathbf{x}} \right] = -\Re \left[\left\{ \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right) \mathbf{p} \right\}^{\dagger} (\mathbf{A}^{-1})^{\dagger} \mathbf{R}_d^t \Delta \mathbf{d} \right] \quad (1.54)$$

$$= -\Re \left[\mathbf{p}^{\dagger} \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right)^{\dagger} (\mathbf{A}^{-1})^{\dagger} \mathbf{R}_d^t \Delta \mathbf{d} \right] = -\Re \left[\mathbf{p}^t \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right)^t (\mathbf{A}^{-1})^t \mathbf{R}_d^t \Delta \mathbf{d}^* \right]. \quad (1.55)$$

En écrivant le problème direct $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, où \mathbf{I} est la matrice identité, on se rend compte que l'opérateur \mathbf{A}^{-1} est l'inverse du problème direct, et représente les fonctions de Green pour des sources impulsionnelles émises en chacun des points du milieu. Pour les systèmes self-adjoints (du second-ordre notamment), la réciprocité spatiale des fonctions de Green (Claerbout, 1976; Aki & Richards, 2002) permet d'écrire la relation $(\mathbf{A}^{-1})^t = \mathbf{A}^{-1}$, ce qui donne l'expression finale du gradient au point m_i

$$\mathcal{G}_{m_i} = \frac{\partial \mathcal{C}(m_i)}{\partial \mathbf{m}} = -\Re \left[\mathbf{p}^t \left(\frac{\partial \mathbf{A}}{\partial m_i} \right)^t \mathbf{A}^{-1} \mathbf{R}_d^t \Delta \mathbf{d}^* \right], \quad (1.56)$$

où t indique la transposée d'une matrice. La matrice dite de rayonnement $(\partial \mathbf{A} / \partial m_i)$ est obtenue par différentiation de \mathbf{A} et est extrêmement creuse, les coefficients non nuls étant associés au point diffractant et à ses voisins. Elle représente la manière dont rayonne le champ diffracté par un paramètre à la position m_i en fonction de l'angle d'ouverture, comme illustré par le

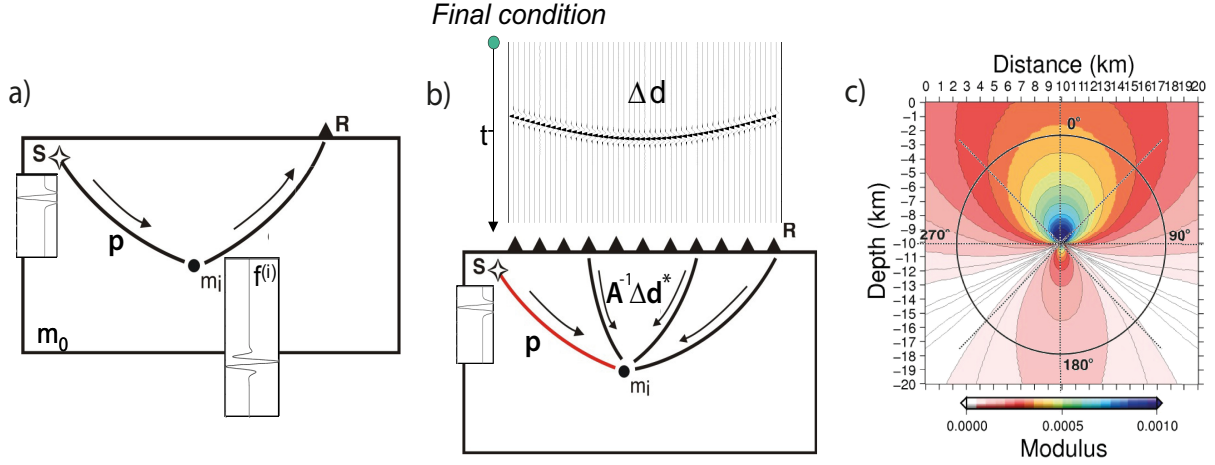


FIGURE 1.9 – Le gradient peut être vu comme (a) la corrélation au niveau des récepteurs du champ diffracté par une source fictive localisée au point m_i et du résidu des données $\Delta \mathbf{d}$; ou bien (b) comme la corrélation au point diffractant m_i entre un champ incident \mathbf{p} et le champ rétropropagé ($\mathbf{A}^{-1}\Delta \mathbf{d}^*$) depuis les récepteurs. (c) Diagramme de rayonnement associé à un diffractant de densité, rayonnant essentiellement aux courts angles d'ouverture. Figures extraites de Operto (2006).

diagramme de rayonnement de la densité dans la Figure 1.9c. Le rayonnement des paramètres dépend de la paramétrisation de l'équation d'onde adoptée, et l'étude des rayonnements permet de mieux comprendre les couplages se produisant entre les paramètres (Tarantola, 1986; Forgues, 1996). L'équation 1.56 indique que le gradient est formé par la corrélation du champ incident \mathbf{p} et du champ adjoint rétropropagé $\mathbf{A}^{-1}\Delta \mathbf{d}^*$ (voir Figure 1.9b), et pondéré par la matrice de rayonnement. On peut faire un lien entre cette formulation et le principe d'imagerie de Claerbout (1985) utilisé en migration, dictant qu'un réflecteur existe si le champ issu de la source et celui issu du récepteur coïncident en temps et en espace. La différence entre la FWI et la migration réside dans la nature du champ rétropropagé. Celui-ci est formé par les résidus des données en FWI alors que ce sont les données elles-mêmes (prétraitées pour enlever tout ce qui n'est pas issu d'une réflexion simple) que l'on considère en migration. Cette formulation peut aussi être reliée à l'approximation de Born, car le gradient évalué au point m_i est uniquement fonction de la corrélation des champs au point m_i , et ne rend donc pas compte des diffractions doubles. En vertu du principe de Huygens, le gradient final est formé par la somme des contributions élémentaires associées à chaque point diffractant, de sorte que $\mathcal{G} = \sum_{i=1}^N \mathcal{G}_{m_i}$, où $N = n_x \times n_z$.

Une alternative à la méthode de l'état adjoint pour calculer le gradient de la fonction coût \mathcal{C} , est de calculer explicitement les dérivées de Fréchet $\mathbf{J}_k = \partial \mathbf{p}(\mathbf{m}_k) / \partial \mathbf{m}$ à chaque itération k , le gradient de \mathcal{C} s'exprimant comme $\mathcal{G}_{m_k} = \mathbf{J}_k^t \Delta \mathbf{d}^*$. Pour obtenir l'expression des dérivées de Fréchet, considérons que la variable d'état $\mathbf{p} + \delta \mathbf{p}$ soit une réalisation physique associée au modèle $\mathbf{m} + \delta \mathbf{m}$, de sorte que : $F(\mathbf{p} + \delta \mathbf{p}, \mathbf{m} + \delta \mathbf{m}) = 0$. Un développement de Taylor au premier ordre de cette nouvelle équation d'état donne

$$F(\mathbf{p} + \delta \mathbf{p}, \mathbf{m} + \delta \mathbf{m}) = F(\mathbf{p}, \mathbf{m}) + \frac{\partial F(\mathbf{p}, \mathbf{m})}{\partial \tilde{\mathbf{p}}} \delta \mathbf{p} + \frac{\partial F(\mathbf{p}, \mathbf{m})}{\partial \mathbf{m}} \delta \mathbf{m} = 0. \quad (1.57)$$

Le terme $F(\mathbf{p}, \mathbf{m})$ étant nul, nous obtenons l'égalité

$$\frac{\partial F(\mathbf{p}, \mathbf{m})}{\partial \tilde{\mathbf{p}}} \delta \mathbf{p} = - \frac{\partial F(\mathbf{p}, \mathbf{m})}{\partial \mathbf{m}} \delta \mathbf{m}. \quad (1.58)$$

La définition de F (équation 1.47) permet d'écrire

$$\mathbf{A} \delta \mathbf{p} = - \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right) \mathbf{p} \delta \mathbf{m} \quad \text{ou encore, au point } m_i \quad \mathbf{A} \frac{\partial \mathbf{p}}{\partial m_i} = - \left(\frac{\partial \mathbf{A}}{\partial m_i} \right) \mathbf{p}. \quad (1.59)$$

On constate que la dérivée partielle du champ \mathbf{p} par rapport au paramètre m_i est le champ propagé par une source fictive donnée par $(\partial \mathbf{A} / \partial m_i) \mathbf{p}$. Cette source est centrée spatialement sur la position de m_i et temporellement sur le temps d'arrivée du champ incident \mathbf{p} au point m_i (voir Figure 1.9a). C'est ainsi que l'on fait l'interprétation de $(\partial \mathbf{p} / \partial m_i)$ en terme de champ diffracté par m_i (Pratt *et al.*, 1998).

Le gradient de la fonction coût est formé par la corrélation à décalage nul entre le champ dérivée partielle $\partial \mathbf{p} / \partial m_i$ et les résidus. Quand ces deux champs sont en phase, une perturbation significative est ajoutée à la position m_i . Le champ dérivée partielle étant associé au champ émis par le diffracteur 'i', les résidus peuvent donc être interprétés comme l'enregistrement du champ d'onde diffracté par les hétérogénéités manquantes, justifiant ainsi l'appellation de tomographie en diffraction.

La comparaison des équations 1.52 et 1.59 permet de mieux cerner les implications numériques du calcul du gradient par la méthode de l'état adjoint, ou par l'approche des dérivées de Fréchet. Dans les deux cas, l'obtention du gradient nécessite une factorisation LU, et la résolution de N_S substitutions associées à l'équation $\mathbf{A} \mathbf{p} = \mathbf{s}$ (avec N_S le nombre de sources). Suite à cela, il faut résoudre dans l'approche par dérivées de Fréchet une substitution par point du milieu ($n_x \times n_z$), le second membre de l'équation 1.59 dépendant de m_i ; alors que la méthode de l'état adjoint requiert N_S substitutions associées à la rétropropagation des résidus par point de tir (équation 1.52), ce qui s'avère être beaucoup plus économique (Pratt *et al.*, 1998).

D'un point de vue de l'implémentation, la construction du gradient (équation 1.56) réalisée dans le code GD de Brossier (2011b) comporte quelques spécificités. Les champs solutions d'ordre P_k , pour $k > 0$, sont projetés sur le barycentre de la cellule, de sorte que la discrétisation des propriétés physiques et des solutions soient identiques, et consistentes avec une interpolation P_0 . La matrice de rayonnement est de même construite suivant une interpolation P_0 . Cette stratégie permet un gain en mémoire et en temps d'assemblage du gradient d'un facteur 3 et 6, suivant que l'ordre d'interpolation pour la modélisation des champs soit de 1 ou de 2 (à cause du plus faible nombre de degrés de libertés supplémentaires en P_0). Par ailleurs, alors que la modélisation des champs incidents et rétropropagés se fait en résolvant un système d'ordre 1 en vitesses/contraintes tels que ceux écrits en 1.5 ou 1.8, le calcul du gradient est réalisé en considérant un système d'ordre 2 en vitesses. Pour cela, les champs auxiliaires de contraintes sont éliminés et seuls les champs de vitesses sont considérés. La matrice de rayonnement est par conséquent aussi calculée en adoptant un schéma de discrétisation associé au système d'ordre 2. Cette approche permet de réduire la mémoire d'un facteur 2.5 (2 champs à stocker au lieu de 5), et présente surtout l'intérêt de rendre le système self-adjoint (Gholami *et al.*, 2012), grâce auquel on peut utiliser le même opérateur pour le calcul du champ incident et du champ adjoint. Ces étapes de projection en P_0 et de restriction aux vitesses nécessitent au préalable une distribution optimisée sur les différents processeurs des solutions obtenues après substitution, grâce à l'algorithme METIS (Karypis & Kumar, 1999), afin d'être sûr que l'ensemble des

solutions associées à un degré de liberté d'une cellule soient sur le même processeur (Brossier, 2011b).

1.2.3.3 Construction et interprétation du Hessien

Pour interpréter le deuxième terme du Hessien (équation 1.26), fonction de la dérivée seconde des données par rapport aux paramètres, nous pouvons dériver l'égalité 1.59 par rapport à m_j . Nous obtenons

$$\mathbf{A} \frac{\partial \mathbf{p}}{\partial m_i \partial m_j} = - \left[\frac{\partial \mathbf{A}}{\partial m_i} \frac{\partial \mathbf{p}}{\partial m_j} + \frac{\partial \mathbf{A}}{\partial m_j} \frac{\partial \mathbf{p}}{\partial m_i} + \frac{\partial^2 \mathbf{A}}{\partial m_i \partial m_j} \mathbf{p} \right]. \quad (1.60)$$

Le membre de droite est composé de trois termes de source fictive. Le premier peut être interprété comme une source localisée en x_i , excitée par le champ aux dérivées partielles $\partial \mathbf{p} / \partial m_j$. Ce champ aux dérivées partielles étant lui-même un champ diffracté par m_j , le premier terme correspond à la diffraction double passant par le point x_j puis par le point x_i . De la même manière, le deuxième terme de source correspond à la double diffraction d'un champ par les points x_i puis x_j . Le troisième terme dépend du diagramme de rayonnement du paramètre (et donc de la paramétrisation), et est généralement nul pour i différent de j (Pratt *et al.*, 1998). Ainsi le deuxième terme du Hessien est principalement associé aux diffractions doubles. Ce terme est ignoré pour les problèmes quasi-linéaires, ainsi que pour les problèmes de grandes tailles en raison d'un coût numérique considérable. La prise en compte de ce terme peut devenir nécessaire lors de l'inversion des hautes fréquences, car le champ d'onde est sensible à de plus petites hétérogénéités.

Le Hessien approximé \mathbf{H}_A , soit la corrélation des dérivées de Fréchet, peut se réécrire en utilisant l'expression 1.59, comme

$$\mathbf{H}_{A, m_i, m_j} = \Re \left[\mathbf{J}_{m_i}^\dagger \mathbf{J}_{m_j} \right] = \Re \left[\mathbf{p}^\dagger \left(\frac{\partial \mathbf{A}}{\partial m_i} \right)^\dagger \mathbf{A}^{-1} \mathbf{R}_d^t \mathbf{R}_d \mathbf{A}^{-1} \left(\frac{\partial \mathbf{A}}{\partial m_j} \right) \mathbf{p} \right], \quad (1.61)$$

où \mathbf{R}_d est l'opérateur de restriction aux récepteurs. Les termes diagonaux du Hessien approximé sont formés par l'auto-corrélation des champs aux dérivées partielles, et contiennent donc le carré de l'amplitude de ces champs. Ces champs aux dérivées partielles étant eux-mêmes fonction du produit des fonctions de Green de la source et du récepteur, la division du gradient (soit $\mathbf{J}^t \Delta \mathbf{d}^*$) par la diagonale du Hessien approximé permet d'éliminer les effets de décroissance géométriques des fonctions de Green de la source et du récepteur. Pour une acquisition de surface, le préconditionnement du gradient par la diagonale du Hessien approximé permet de mettre à l'échelle les perturbations profondes du modèle, contraintes par des traces à grands offsets et de faibles amplitudes, relativement aux perturbations superficielles contraintes par des traces courts offsets et de larges amplitudes. Ce préconditionnement par la diagonale du Hessien approximé est judicieux en raison de la structure à diagonale dominante du Hessien (Pratt *et al.*, 1998), et car nous voulons obtenir son inverse. Si N_S et N_R sont le nombre de sources et de récepteurs, le nombre de phases de substitutions est de N_S pour calculer les N_S \mathbf{p}_i (obtenus lors du calcul du gradient), et de N_R pour résoudre le problème $\mathbf{A} (\mathbf{R}_d \mathbf{A}^{-1}) = \mathbf{I}$, où $(\mathbf{R}_d \mathbf{A}^{-1})$ est formé des N_R colonnes associées à chaque récepteur, et chaque colonne de \mathbf{I} comprend une seule valeur non nulle de 1, à la ligne associée à la position du récepteur considéré. Le calcul des dérivées de Fréchet aux récepteurs $(\mathbf{R}_d \mathbf{A}^{-1} (\partial \mathbf{A} / \partial m_j) \mathbf{p})$ nécessite donc $N_S + N_R$ phases

de substitutions (Shin *et al.*, 2001; Virieux & Operto, 2009). A noter qu'il fut proposé par Shin *et al.* (2001) de calculer un pseudo-Hessien approximé (diagonal) pour lequel les fonctions de Green associées aux récepteurs ne sont pas prises en compte, n'entraînant donc pas de surcoût par rapport au calcul du gradient. Nous comparons dans la partie 3.2.6 les conséquences de l'utilisation de la diagonale du Hessien approximé et du pseudo-Hessien approximé sur notre application au champ de Valhall.

Dans un cas idéal, un point diffractant porte une information sur l'ensemble des nombres d'ondes, ce qui est possible d'après l'expression 1.63 si la source est un dirac (source impulsionnelle), et/ou si l'éclairage est parfait. En réalité, la source est à bande passante limitée, et un dispositif de surface est en général utilisé. Les termes non-diagonaux du Hessien approximé, formés par l'inter-corrélation des champs aux dérivées partielles des différents paramètres, jouent le rôle d'opérateur de déconvolution de l'image tomographique, afin de mieux séparer la contribution de chaque point diffractant, et donc de *déflouter* l'image. Tout comme en traitement du signal, l'aplatissement du spectre des nombres d'ondes par déconvolution a aussi pour effet d'augmenter la part de bruit (haute fréquence) dans l'image. Un lissage Gaussien ($G_{\mathbf{m}}$, équation 1.62) est alors appliqué au modèle de perturbation. Il est aussi ajouté au Hessien un facteur *d'amortissement* de l'inversion (ϵ , dans l'équation 1.62), afin d'une part de prévenir les divisions par zéro, et d'autre part de rendre la matrice Hessienne mieux conditionnée en omettant les valeurs propres les plus faibles. Le choix de cette variable est crucial, une valeur trop petite aboutissant à des instabilités, et une valeur trop grande peut complètement annihiler l'action du Hessien approximé. Nous comparons les résultats obtenus pour deux valeurs de ϵ dans la partie 3.2.6. Dans le cas où la fonctionnelle est uniquement fonction d'un terme de résidu sur les données (équation 1.21), et où l'on préconditionne le gradient par la diagonale du Hessien, la perturbation de modèle prend la forme

$$\Delta \mathbf{m} = -\alpha [\text{diag } \mathbf{H} + \epsilon \text{Max}(\text{diag } \mathbf{H})]^{-1} G_{\mathbf{m}} Re \left[\mathbf{p}^t \left(\frac{\partial \mathbf{A}^t}{\partial m_i} \right) \mathbf{A}^{-1} \mathbf{R}_d^t \Delta \mathbf{d}^* \right], \quad (1.62)$$

où α est le pas de descente, et $\text{diag}(\mathbf{H})$ correspond à la diagonale du \mathbf{H}_A ou du \mathbf{H}_S .

Par ailleurs, le Hessien joue un rôle clé dans le cadre de l'inversion multiparamètre. Il permet d'une part de mettre à l'échelle les différentes classes de paramètres, notamment celles ayant une faible empreinte dans les données. Pour ce type de paramètre, les dérivées de Fréchet sont de très faibles amplitudes. La division d'un gradient de faible amplitude ($\mathbf{J}^t \Delta \mathbf{d}^*$) par un Hessien d'encore plus petite amplitude ($\mathbf{J}^t \mathbf{J}$) est censée redonner au paramètre secondaire sa véritable amplitude. D'autre part, les blocs hors-diagonaux du Hessien, formé des inter-corrélations des dérivées de Fréchet associées à deux classes de paramètres différentes, sont censés corriger des effets de couplages (ou sensibilités croisées) des paramètres.

1.2.3.4 Analyse de résolution

Une analyse de résolution détaillée des méthodes de tomographie en diffraction, dont l'inversion des formes d'ondes fait partie, a été développée par Devaney (1982); Miller *et al.* (1987); Wu & Toksöz (1987); Sirgue & Pratt (2004). Le nombre d'onde \mathbf{k} imagé à proximité d'un point diffractant est relié à la fréquence f de la source et à l'ouverture angulaire θ par :

$$\mathbf{k} = \frac{2f}{c} \cos \frac{\theta}{2} \bar{\mathbf{q}}, \quad (1.63)$$

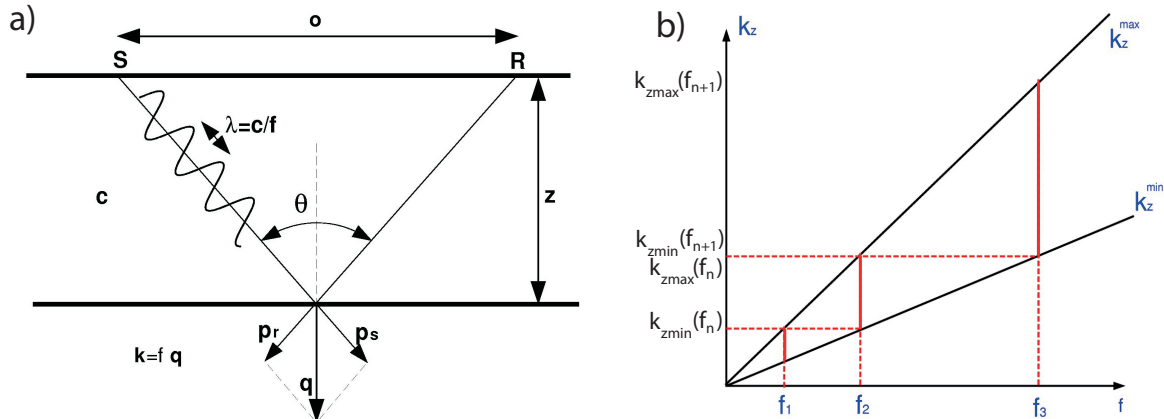


FIGURE 1.10 – (a) Illustration schématique de la relation entre les paramètres de l'expérience sismique et la résolution de l'imagerie sismique dans le cas particulier d'une onde réfléchie sur un interface plan (voir équation 1.63). (b) Illustration de la stratégie de discrétisation en fréquence. Une seule fréquence image une gamme de nombre d'onde (Sirgue & Pratt, 2004).

où \mathbf{k} est le nombre d'onde du modèle à imager, θ l'ouverture angulaire entre une source, un point diffractant et un capteur ; et c la vitesse de l'onde traversant le milieu à imager. Je précise que \mathbf{k} n'est pas un nombre d'onde au sens classique ($= \frac{2\pi}{\lambda}$), mais représente une fréquence spatiale. Le vecteur unitaire $\bar{\mathbf{q}}$ est orienté dans la direction du vecteur lenteur \mathbf{q} . Ces différents paramètres sont illustrés dans le cas particulier d'une onde réfléchie sur une interface plane situé à la base d'un demi-espace homogène sur la Figure 1.10a.

L'expression (1.63) montre tout d'abord qu'une ouverture angulaire et une fréquence dans l'espace des données imagent un nombre d'onde dans l'espace des modèles. Ces deux paramètres contrôlent la résolution de l'image sismique, au sens où plusieurs couples (f, θ) imagent un même nombre d'onde. D'autre part, cette formule confirme que plus l'ouverture devient grande, plus l'imagerie est sensible aux grandes longueurs d'onde. Ces observations sont illustrées sur la Figure 1.11, où l'on montre en haut à gauche une onde plane incidente et au centre une onde plane diffractée. Le panneau de droite est obtenu par multiplication des deux ondes planes, et représente l'image monochromatique du milieu si celui-ci n'est éclairé que par une seule fréquence (5 Hz) et une seule ouverture angulaire (10°). Dans la deuxième série de panneau, la fréquence est maintenue constante mais l'ouverture a été augmentée pour atteindre 120° . Le nombre d'onde imagé devient alors plus faible. La série de panneau du bas montre qu'avec une fréquence plus élevée (10 Hz) et une même ouverture angulaire que le panneau du milieu, on retrouve le même nombre d'onde que le panneau du haut, illustrant la redondance d'information fournie par le double contrôle fréquence-ouverture.

Cette redondance d'information permet de n'inverser que quelques fréquences lors de l'inversion. En effet, une seule fréquence image une gamme de nombre d'onde d'autant plus importante que l'ouverture angulaire est grande. La relation 1.63 montre que les nombres d'onde minimum et maximum varient linéairement avec la fréquence (Figure 1.10b) (Sirgue, 2003). La bande de nombre d'onde imagée à une fréquence donnée est plus étroite que celle imagée par une fréquence supérieure, d'où la nécessité d'échantillonner plus finement les basses fréquences. L'intervalle des fréquences à inverser doit être suffisamment faible pour garantir l'échantillon-

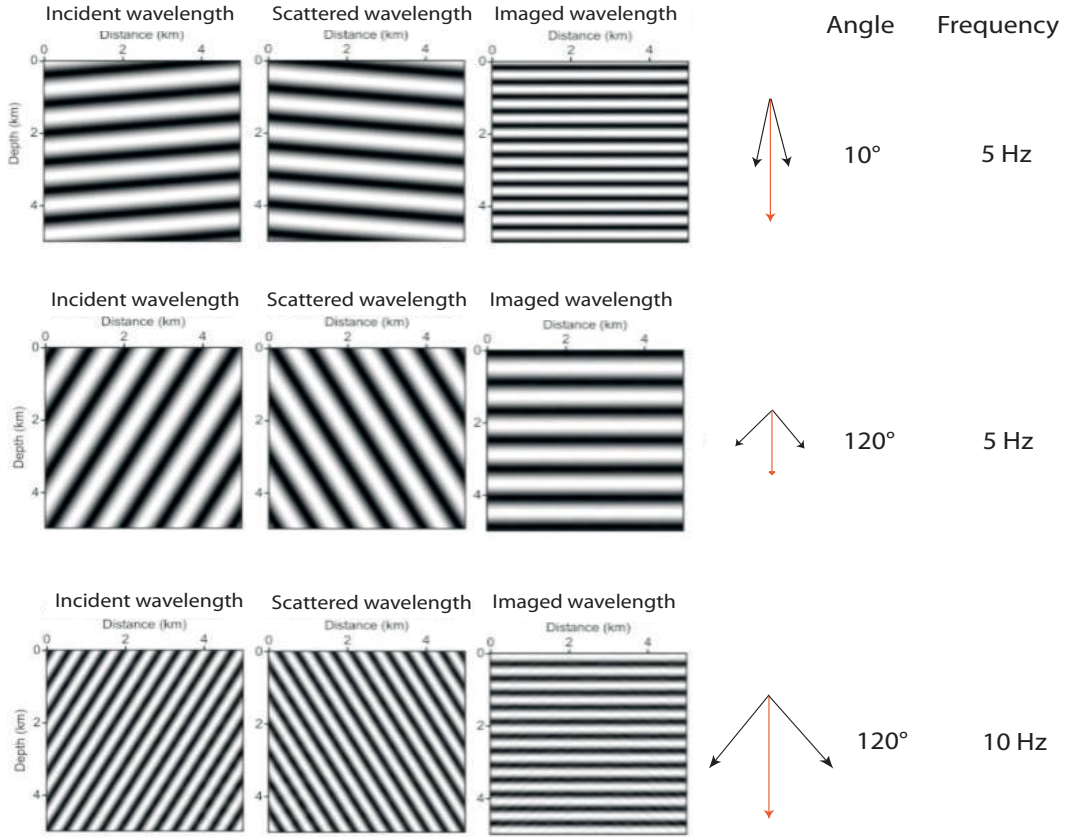


FIGURE 1.11 – Illustration du double contrôle exercé par la fréquence et l'ouverture angulaire sur l'amplitude et la phase du nombre d'onde imagé. Adapté de Sirgue (2003).

nage continu des nombres d'ondes, mais assez large pour éviter le calcul redondant d'un nombre d'onde.

1.2.3.5 Approches multi-échelles et algorithmes

L'algorithme utilisé dans les différents codes de FWI est structuré en quatre boucles principales emboîtées (Figure 1.12). La boucle externe parcourt les groupes de fréquences inversés de manière croissante, afin de réduire la non-linéarité de l'inversion par la reconstruction des grandes longueurs d'ondes spatiales du milieu avant les plus courtes. Un groupe de fréquences est un ensemble de fréquences adjacentes inversées simultanément, afin de tirer bénéfice de la redondance des informations sur les nombres d'ondes spatiaux, et ainsi améliorer le rapport signal sur bruit de l'imagerie. Dans la plupart des tests présentés dans la thèse, nous inversons trois fréquences par groupes de largeur 0.5 Hz environ, avec un recouvrement entre la dernière et la première fréquence de deux groupes successifs. La largeur de bande d'un groupe de fréquence doit être assez fine pour empêcher les artéfacts liés aux sauts de phase hautes fréquences, et assez large pour inverser simultanément assez de fréquences pour stabiliser l'inversion, tout en évitant de suréchantillonner inutilement les fréquences pour gagner en temps de simulation (Brossier *et al.*, 2009b).

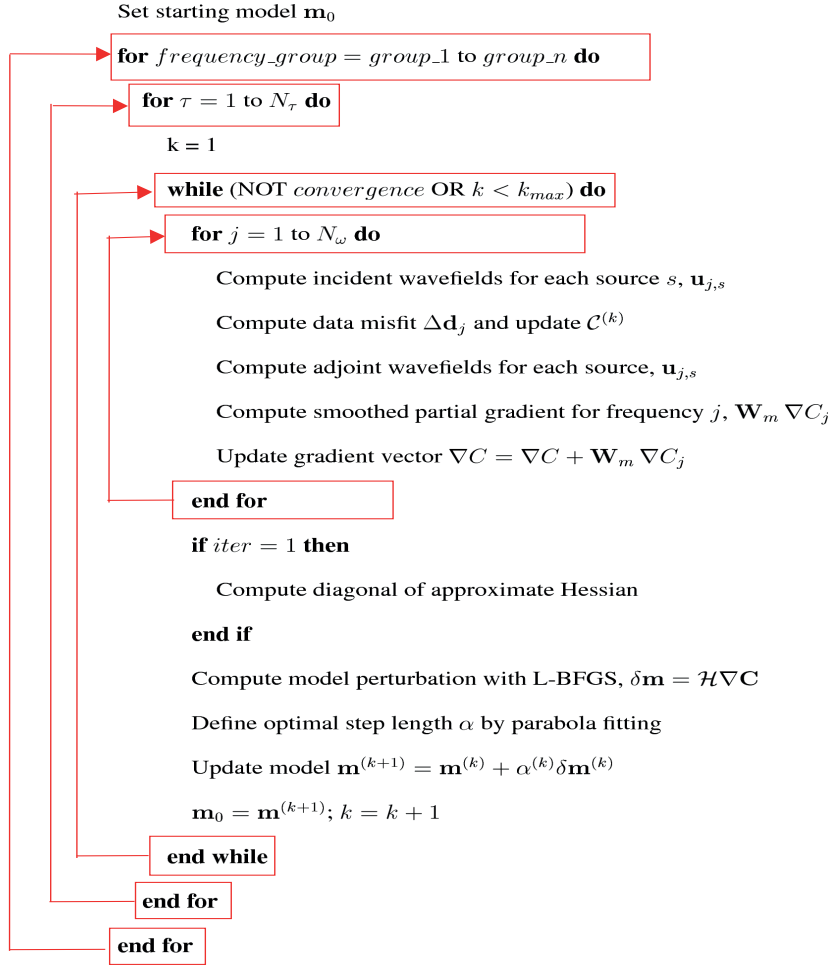


FIGURE 1.12 – Algorithme d'inversion des formes d'onde en fréquence.

La deuxième boucle porte sur la partie imaginaire des fréquences, associées à l'amortissement temporel des données (Sirgue, 2003; Brenders & Pratt, 2007b; Brossier *et al.*, 2009a; Shin & Cha, 2009). En effet, si $f(t)$ est le champ d'onde temporel et $F(\omega)$ le champ d'onde dans le domaine des fréquences, l'amortissement exponentiel en temps équivaut à ajouter un argument complexe à $F(\omega)$, tel que

$$F\left(\omega + \frac{i}{\tau}\right) \exp \frac{t_0}{\tau} = \int_{-\infty}^{+\infty} f(t) \exp^{-\frac{t-t_0}{\tau}} \exp^{-i\omega t} dt, \quad (1.64)$$

où τ est le paramètre d'amortissement exponentiel des données, et t_0 le temps de première arrivée, utilisé ici comme origine temporelle de la fonction exponentielle décroissante. L'utilisation de fréquences complexes permet de rajouter un niveau hiérarchique supplémentaire dans l'approche multi-échelle (en plus de celui permis par la fréquence), en privilégiant au début de l'inversion les premières arrivées, puis en considérant dans la suite de l'inversion des données contenant des arrivées plus tardives. Les premières arrivées correspondent en effet aux grandes ouvertures angulaires qui permettent d'imager les bas nombres d'ondes du modèle, rendant ainsi le problème plus linéaire. Une procédure équivalente est utilisée pour les algorithmes de

FWI basés sur une modélisation en temps, basé sur un fenêtrage (plus) flexible en offset et en temps des données. Notons que l'équivalent d'un fenêtrage en temps n'est pas réalisable en fréquence, car il correspond à une convolution en fréquence, ce qui n'est pas possible car toutes les fréquences ne sont pas disponibles en même temps. Nous montrons sur la Figure 1.13 l'effet d'un amortissement temporel exponentiel à partir de la première arrivée sur des sismogrammes.

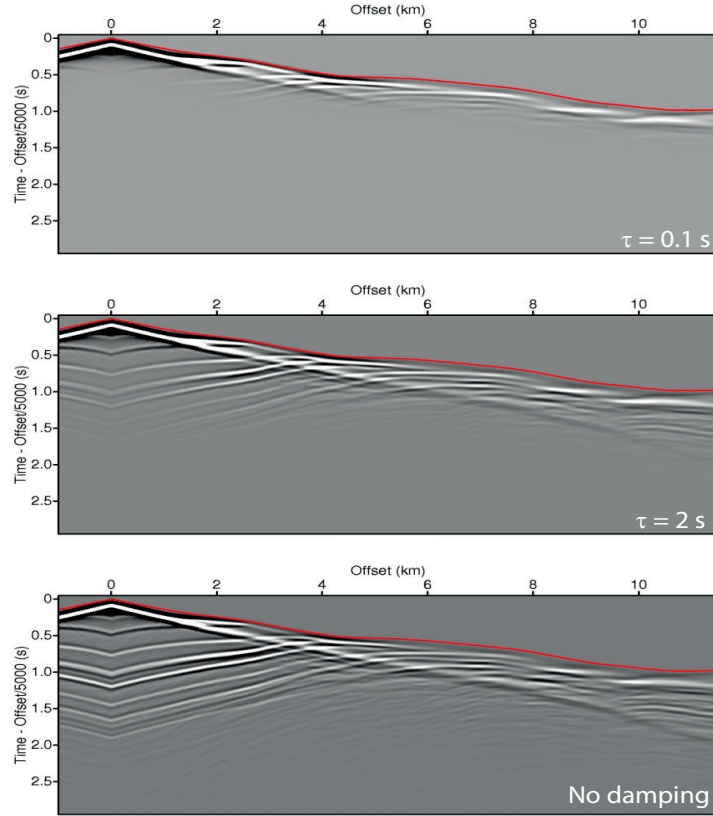


FIGURE 1.13 – Comparaison des sismogrammes synthétiques calculés sans amortissement (c), et avec un paramètre d'amortissement τ de (a) 0.1 s, et (b) 2 s, appliqué à partir de la première arrivée dénotée par la ligne rouge. Extrait de Operto (2006).

La troisième boucle parcourt les itérations de l'inversion du groupe de fréquences considéré, afin d'assurer la convergence vers le minimum de la fonction coût. Un modèle est créé à chaque itération de l'inversion d'un groupe de fréquence, pour être réinjecté comme modèle initial de l'itération suivante. La dernière boucle parcourt les fréquences appartenant au groupe de fréquence considéré.

1.2.3.6 Régularisations

Nous avons jusqu'à présent développé les équations normales et la perturbation de modèle associées à une fonctionnelle ne dépendant que du résidu des données (équations 1.21, 1.25 1.26 et 1.62). Le problème inverse étant par nature très non-linéaire, il est nécessaire de rajouter un terme de contrainte sur les modèles dans la fonction coût afin de stabiliser l'inversion et de

restreindre la solution autour d'un modèle *a priori*. Celui-ci peut être un modèle initial lisse décrivant la cinématique, et/ou un modèle construit à partir d'a priori géologiques tels que ceux fournis par des logs de puit. Nous verrons notamment dans la partie 3.3.4 que l'apport de la *régularisation Tikhonov*, développée ci-dessous, permet de stabiliser l'inversion lorsque le L-BFGS est utilisé.

Dans le cadre de la régularisation Tikhonov, la fonctionnelle des données est augmentée par des termes de régularisations dans l'espace des modèles, tel que

$$C(m) = \frac{1}{2} \sum_{j=1}^{N_C} \Delta \mathbf{d}_j^\dagger \mathbf{W}_{d_j} \Delta \mathbf{d}_j + \frac{1}{2} \sum_{i=1}^{N_P} \lambda_i (\mathbf{m}_i - \mathbf{m}_{prior_i})^\dagger \mathbf{W}_{m_i} (\mathbf{m}_i - \mathbf{m}_{prior_i}), \quad (1.65)$$

où N_P et N_C sont le nombre de classes de paramètres et le nombre de composantes des données respectivement, et où \mathbf{W}_d et \mathbf{W}_m sont les matrices de pondération des données et des modèles. Les termes de données et de modèles sont pondérés par les scalaires d'amortissement de l'inversion λ_i . Le modèle *a priori* (\mathbf{m}_{prior}) est pris comme le dernier modèle trouvé à la fréquence précédente. Un gain avec l'offset, ou un mute des arrivées inférieures à un certain offset peuvent être implémentés par l'intermédiaire de \mathbf{W}_d . \mathbf{W}_m , l'inverse de la matrice de covariance, est composé d'un terme de lissage et d'un terme d'amortissement σ . Celui-ci exprime l'incertitude du paramètre, il est par conséquent fonction de l'ordre de grandeur du paramètre, et rend ainsi la fonction coût de l'espace des modèles adimensionnelle. L'expression de l'inverse de \mathbf{W}_m est

$$\mathbf{W}_{m_i}^{-1}(z, x, z', x') = \sigma_i^2(z, x) \exp\left(\frac{-|x - x'|}{x_{cor}(z, x)}\right) \exp\left(\frac{-|z - z'|}{z_{cor}(z, x)}\right), \quad (1.66)$$

où $|\cdot|$ est une valeur absolue, x_{cor} et z_{cor} sont les longueurs de corrélations du lissage, définies comme une fraction de la longueur d'onde. L'opérateur de lissage \mathbf{W}_m^{-1} est fonction d'une exponentielle dont l'argument est une valeur absolue, afin d'obtenir analytiquement l'inverse (\mathbf{W}_m) de cet opérateur dont l'expression est nécessaire dans la fonction coût. \mathbf{W}_m est alors une matrice tridiagonale (Tarantola (1987), paragraphe 7.21 pages 308-309). Plus de détails sur la manière dont fut configurée cette régularisation pour l'appliquer aux données réelles de Valhall sont donnés dans la partie 3.3.4.

La perturbation de modèle de la classe de paramètre i à l'itération k , associée à la fonction coût 1.65 s'écrit

$$\Delta \mathbf{m}_{i,k} = \Re [\mathbf{W}_{m_i}^{-1} \mathbf{J}_{k,i}^t \mathbf{W}_d \mathbf{J}_{k,i}^* + \lambda_i \mathbf{I}]^{-1} \Re [\mathbf{W}_{m_i}^{-1} \mathbf{J}_{k,i}^t \mathbf{W}_d \mathbf{R}_d^t \Delta \mathbf{d}_k^* + \lambda_i \mathbf{I}(\mathbf{m}_{i,k} - \mathbf{m}_{i,prior})]. \quad (1.67)$$

Chapitre 2

Le champ pétrolier de Valhall

Sommaire

2.1	Introduction	67
2.2	Contexte géologique et pétrolier	69
2.3	Etat de l’art des études géophysiques effectuées sur Valhall	73
2.4	Présentation du jeu de données OBC	78
2.4.1	Anatomie des données	78
2.4.2	Composante de pression calculée pour le cas synthétique de Valhall . .	80
2.4.3	Données 4C du dispositif LoFS	80

2.1 Introduction

La motivation première de ma thèse est l’application de la méthode d’inversion des formes d’ondes au champ pétrolier de Valhall, localisé en mer du Nord (Fig. 2.1a). Ce champ est exploité par la compagnie BP depuis 1982, et le sera encore probablement jusqu’en 2050. Plus de 500 millions de barils ont été extraits jusqu’à maintenant, et 500 autres le seront probablement. Cela porte à 38 % le facteur de récupération si l’on considère les 2.6 milliards de barils estimés pour cette zone d’exploitation. Le dispositif d’acquisition 3D (Fig. 2.1b), la répétabilité des acquisitions sismiques tous les six mois, et le panel de méthodes géophysiques et géologiques employées à son étude en font un site de référence exceptionnellement bien contraint (voir la densité de puits, Figure 2.1b). Les applications de la méthode d’inversion des formes d’ondes sur données réelles n’étant pas nombreuses au début de ma thèse, il s’est avéré intéressant de considérer une acquisition bidimensionnelle pour explorer l’influence des différents paramètres de cette méthode d’imagerie.

La zone d’étude à imager en deux dimensions est un profil d’une vingtaine de kilomètres de long sur environ 5 km de profondeur, soit des dimensions caractéristiques d’une cible pétrolière. Comme pour toutes méthodes de prospection géophysique considérant une acquisition de surface, l’image obtenue sera mieux échantillonnée au centre de l’acquisition où la couverture est maximum, elle présentera une dégradation de la résolution avec la profondeur, et elle aura tendance à reconstruire la réflectivité du milieu en profondeur en raison de l’éclairage de cette zone par des données majoritairement associées à des petits angles d’ouvertures.

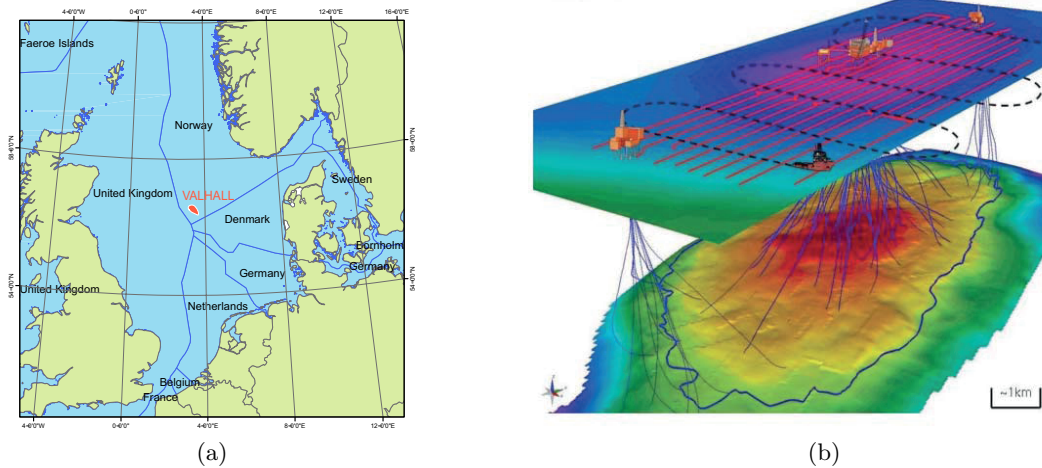


FIGURE 2.1 – (a) Localisation de Valhall en mer du Nord, à la limite des eaux territoriales Norvégiennes, au niveau de la zone 31N du référentiel UTM European Datum 1950. (b) Aperçu général du champ pétrolier de Valhall montrant la localisation des puits de forages (lignes bleues) et des lignes sismiques 4 composantes (lignes rouges) alignées le long de l'anticlinal, dont la surface est imagée par un dégradé de couleur à environ 3 km de profondeur (extrait de van Gestel *et al.* (2008b)).

Pour illustrer le type de milieu que nous aurons à reconstruire, je présente sur la Figure 2.2a le modèle synthétique de Valhall, fourni par BP, et dont les structures sont représentatives de la géologie du vrai milieu. Ce modèle fut construit par O'Brien *et al.* (1999), en se basant sur des études de sismique de surface et de type VSP¹, ainsi qu'à partir de logs soniques, dipole soniques, de densité, de lithologie et de pression. Les modèles synthétiques de vitesse et de densité nous ont été fournis avec un pas de 3.125 m. Nos tests de modélisation ne nécessitant pas d'échantillonnage spatial inférieur à 10 m, nous avons avant de les rééchantillonner, filtré les modèles synthétiques pour omettre les composantes spatiales aux longueurs d'ondes inférieures à 10 m. Le modèle de V_P se distingue par la présence d'une pile de couches fortement chargées en gaz dans des sédiments tertiaires, et de deux interfaces hautes vitesses à 3 km et 5 km de profondeur, séparées par une zone de plus basse vitesse. Les fluides du réservoir sont piégés par un anticlinal délimité par une couche haute vitesse à 3 km de profondeur. Bien que la structure du milieu soit relativement simple au premier ordre en raison de son aspect tabulaire, l'imagerie du champ de Valhall présente des difficultés. C'est d'abord un milieu marin avec une faible profondeur d'eau (70 m), impliquant la présence de forts multiples dans les données. Pour illustrer ce point, nous avons représenté sur la Figure 2.2(b-c), les sismogrammes calculés dans le modèle synthétique de Valhall lorsqu'une condition absorbante ou une surface libre sont implémentées le long du bord supérieur du modèle. Nous constatons que les sismogrammes obtenus avec une surface libre comportent une forme d'onde plus complexe au niveau des réflexions, et que l'onde

1. Vertical Seismic Profile : cette technique consiste à placer les récepteurs le long d'un puit, pour qu'ils enregistrent les tirs émis par des sources localisées en surface à l'aplomb du puit.

directe ainsi que les ondes dispersives guidées dans la proche surface (Robertson *et al.*, 1996) sont bien plus fortes dans les profils en tirs communs.

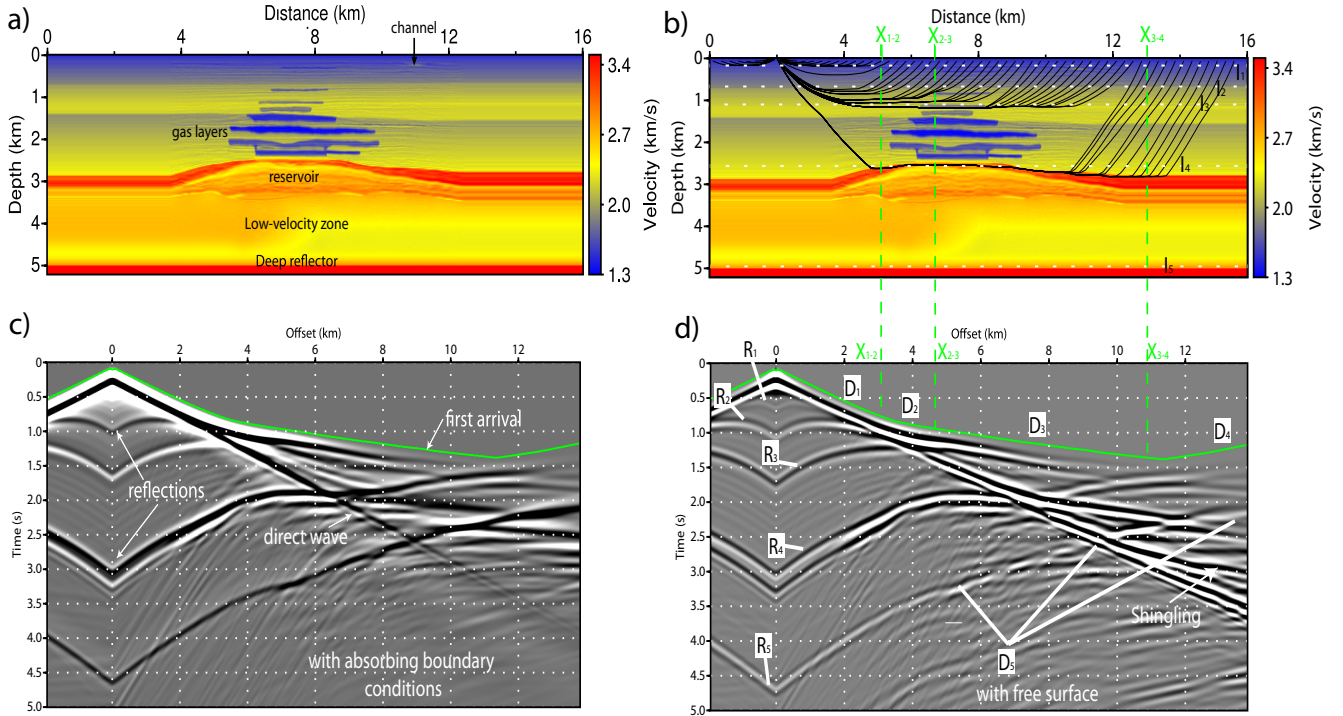


FIGURE 2.2 – (a-b) Modèle synthétique de Valhall. (a) Les principales structures géologiques sont indiquées. (b) Raies de premières arrivées se réfractant le long des 5 principales interfaces I_1 à I_5 . (c-d) Sismogrammes de pression représentés avec une vitesse de réduction de 2.5 km/s, calculés en approximation acoustique dans le modèle en (a), lorsqu’une condition absorbante (c) et une condition de surface libre (d) sont implémentées au toit du modèle. En (d) est rajouté le nom des arrivées réfléchies et réfractées sur les interfaces définies en (b).

Une autre difficulté est causée par la présence de sédiments très fortement chargés en gaz. Ces sédiments présentent de très basses vitesses, ce qui étire vers le bas les images migrées en temps (Figure 2.3) (O. Barkved, 2003). Le gaz atténue de plus les composantes hautes fréquences du signal. D’autre part, l’imagerie sismique des méthodes basées sur les temps de premières arrivées n’est pas adaptée aux milieux comportant des zones de basses vitesses car celles-ci sont sous-échantillonnées.

Enfin, les données de géophones ne sont pas riches en ondes converties de type P-S dans les environnements marins sédimentaires comme celui de Valhall, en raison des faibles contrastes d’impédances dans la proche surface, qui est composée de sédiments faiblement consolidés.

2.2 Contexte géologique et pétrolier

Le contexte géologique est primordial pour mieux comprendre l’origine des structures tectoniques et les différents types de sédiments rencontrés, en vue d’une extraction optimisée des ressources. Pour les explications suivantes, le lecteur pourra se référer à la figure 2.4(a-c), sur

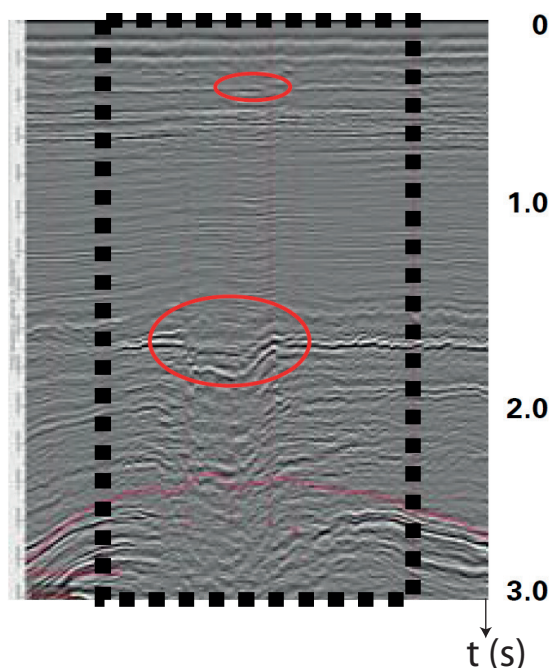


FIGURE 2.3 – Image migrée en temps, étirée verticalement vers le bas (cercles rouges) à cause de la présence de gaz dans les sédiments (extrait de Barkved *et al.* (2010b), figure 1 de l'article).

laquelle est montré un log stratigraphique et une coupe transversale à l'anticlinal, et à la Figure 2.2a pour le modèle synthétique de Valhall, bien que les profondeurs dans ce dernier ne soient pas exactement identiques au cas réel.

Le champ pétrolier de Valhall est un ancien graben ayant joué du Trias jusqu'au Crétacé inférieur (Munns, 1985; Leonard & Munns, 1987). Les roches du Jurassique inférieur et moyen sont absentes des logs stratigraphiques, probablement en raison d'une première surrection de la structure qui a été érodée au Jurassique moyen. Les sédiments argileux du Jurassique supérieur constitueront plus tard la source du réservoir sus-jacent. Le Crétacé inférieur et supérieur correspond à une période de sédimentation lente où le niveau de la mer était anormalement haut, en particulier au centre du graben. Celui-ci s'est comblé de calcaire crayeux (*chalk*) caractérisé par une biomicrite composée de cocolithophoridés et de foraminifères planctoniques. L'origine de ces sédiments fut d'abord d'origine autochtone (par dépôt gravitaire) formant ainsi un premier réservoir nommé *Hod* ; puis d'origine allochtone, c'est-à-dire issue d'un transport de masse via des courants sous-marins, des glissements de terrain ou des courants turbiditiques, formant le second et principal réservoir nommé *Tor*. La raison de cette évolution est à relier au contexte régional de compression (mouvements Laramide), qui s'est poursuivi jusqu'à l'Oligocène. Cette phase a abouti à la formation de la ride de Lindesnes, dont Valhall avec quatre autres champs pétroliers du même type fait partie. Ces mouvements se sont traduits sur Valhall par la surrection d'un anticlinal, dont la croissance a été facilitée par l'inversion des failles préexistantes du graben (Figure 2.4b). Ainsi, le comblement du graben et la surrection de la structure se sont accompagnées d'une diminution progressive de la profondeur d'eau. Cela, conjugué aux mouvements tectoniques de l'époque, et à une alimentation en sédiments par un ancien chenal provenant du nord du champ, a provoqué des glissements de terrain responsables de l'évolution

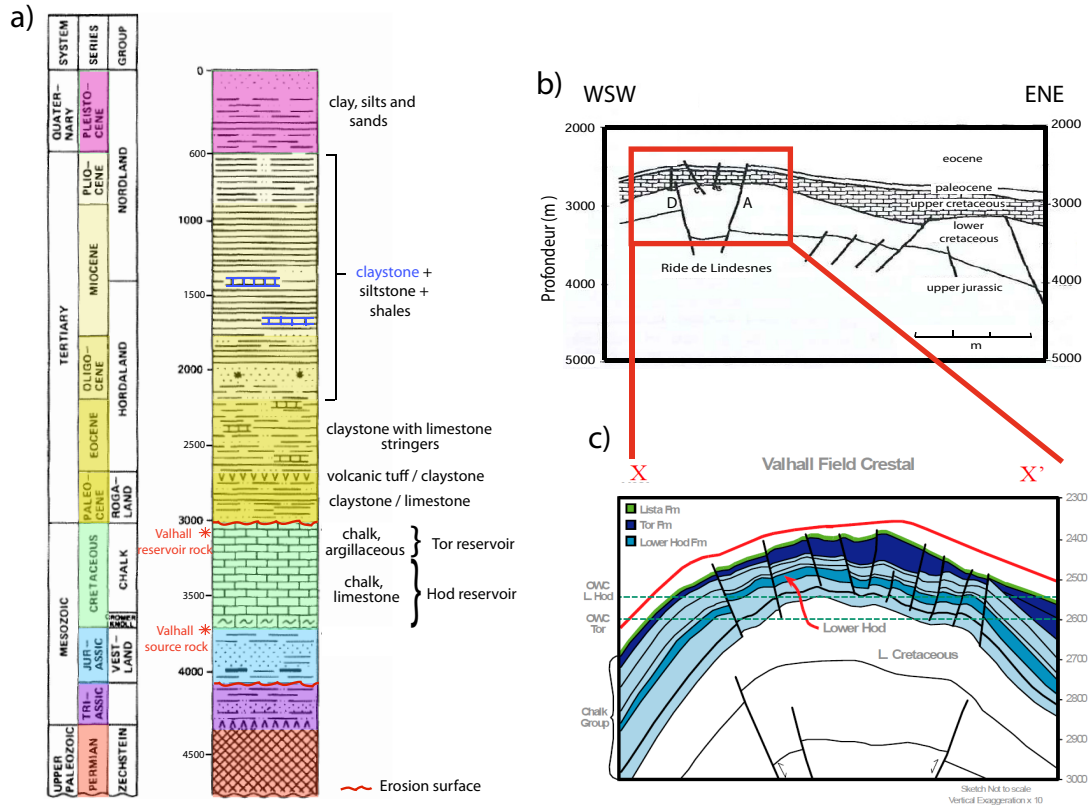


FIGURE 2.4 – (a) Log stratigraphique du champ de Valhall, adapté de Munns (1985) (figure 5 de l'article) et de Olofsson *et al.* (2003). Les profondeurs du log sont plus représentatives des flancs de l'anticlinal que du sommet. (b) Coupe transversale à l'anticlinal (Munns (1985), figure 20 de l'article). Notez la structure caractéristique de graben inversé, la surrection étant rendue possible grâce à l'inversion des failles profondes A et D. On remarquera aussi comment l'épaisseur de la couche du Crétacé supérieur augmente du sommet vers les flancs, indiquant un soulèvement syndépositionnel. Au contraire, la couche du Crétacé inférieur au centre du graben est la plus épaisse car le relief était négatif au centre au moment du dépôt. (c) Agrandissement sur les réservoirs Tor et Hod, compartimentés par une série de failles.

de la sédimentation.

La crête de cet anticlinal a émergé au début du Tertiaire et a subi une phase d'érosion terrestre, donnant lieu à une surface basale tertiaire relativement plate, actuellement à environ 2.6 km de profondeur. Celle-ci recouvre une série de grabens et de semi-grabens dans lesquels des résidus de la formation Tor (du Crétacé supérieur) sont préservés (Figure 2.4c). Les failles des grabens compartimentent le réservoir, ce qui explique que des puits creusés à différents endroits dans le Tor aient relevé des propriétés de porosité, de perméabilité et des pressions de confinement pouvant varier significativement (Mueller *et al.*, 1997). La relative bonne préservation de la formation Tor (50 m d'épaisseur au sommet) est à attribuer à un soulèvement syndépositionnel qui a permis le dépôt d'une importante épaisseur de sédiments, en particulier sur les flancs de la structure.

Au Tertiaire, une forte subsidence associée à un afflux de matériel terrigène a permis de pour-

suivre une sédimentation d'environnement marin peu profond, à l'origine d'une forte accumulation de sédiments boueux (*siltstones*). La nature argileuse de ces sédiments les rend propice à être une source de gaz, bien qu'ils ne soient pas actuellement à un stade optimal de maturation. Ils se sont aussi chargés en gaz car la barrière sommitale de l'anticlinal n'est pas complètement hermétique. Ce gaz est bloqué localement à certaines profondeurs par des couches d'argilite sus-jacentes (*claystones*), notamment au niveau du Miocène et au sommet du Tertiaire (Munns, 1985). Le nuage de gaz présent dans ces sédiments constitue un véritable écran sismique, et est responsable de la plupart des difficultés de forage rencontrées. En effet, la pression dans les poches de gaz étant inférieure à celle des sédiments sus-jacents, la rencontre inattendue de la tête de forage avec cette poche provoque une aspiration de la boue de forage dans les poches si la boue n'a pas une densité appropriée. Le gaz remonte alors à très haute vitesse par le puits et peut provoquer une explosion en surface. Il se peut aussi que, sous la contrainte des roches environnantes, la dépression au sein du puits ne le fasse imploser. Le même problème est rencontré lorsque le puits atteint le réservoir, la pression dans ce dernier étant inférieure à celle des couches sus-jacentes. La boue utilisée sur ce champ est exclusivement non aqueuse ('*oil based*' mud), et non pas aqueuse ('*water based*') en raison de la nature argileuse des sédiments à traverser. Par ailleurs, une nouvelle technologie de puits a dû être utilisée pour faire face aux problèmes de subsidence particulièrement dramatique dans les sédiments du Tertiaire.

Les sédiments quaternaires présents dans les 600 premiers mètres sont composés d'argiles, de boue et de sable, et se sont déposés dans un environnement peu profond (comme actuellement). On y trouve par exemple des anciens chenaux, qui sont la marque laissée par une rivière se jetant dans la mer. On les reconnaît aisément dans la figure 9.b de Sirgue *et al.* (2010), ou encore dans le modèle synthétique de Valhall (à $X = 11$ km).

Les réservoirs Tor et Hod se sont remplis suite à une migration de pétrole issu des roches mères du Jurassique supérieur situées vers 4 km de profondeur. La migration de fluides fut facilitée par la présence de nombreuses fractures créées lors des différentes phases tectoniques, et atteint un pic de nos jours. La formation Tor présente une porosité primaire pouvant atteindre 50% au sommet et 36% sur les flancs. La porosité de la biomicrite la composant était de 80% au moment du dépôt, mais la pression des 3000 m de sédiments devrait être responsable d'une porosité actuelle de 10%. La préservation de cette forte porosité fut possible d'une part grâce à la phase de consolidation mécanique initiale (suite aux mouvements de compression), qui a permis un réarrangement des grains et par conséquent une réduction de l'espace inter-granulaire. D'autre part, le fait que les sédiments du Tor aient été gorgés d'eau jusqu'à la fin du Miocène, puis de pétrole par la suite (jusqu'à 95% de saturation), a aussi pour une large part empêché le phénomène naturel de compaction mécanique. Le pourcentage de porosité et de perméabilité s'effondre rapidement lorsque le matériel se vide de ses fluides, ce qui explique la forte subsidence actuelle : près de 6 m en 30 ans en surface.

La perméabilité dans les réservoirs de Valhall est faible (moins de 10 mDarcy), à cause des grains très fins (1 à 5 μm) du calcaire. En réalité, la perméabilité réelle est un ordre de grandeur supérieure à celle mesurée dans les échantillons grâce à un réseau de fractures très développé, en particulier au sommet de l'anticlinal, dont les techniques d'extraction tirent profit en réalisant de la fracturation hydraulique.

Le pétrole extrait est à un stade précoce de maturité et est de bonne qualité, avec un indice gravitaire API (*American Petroleum Institute*) de 36, qui quantifie le niveau de légèreté du pétrole par rapport à l'eau. Cet indice est calculé par la formule : $API_{liquide} = (141.5\rho_{eau}/\rho_{liquide}) - 131.5$. L'eau a donc un indice API de 10, le pétrole entre 10 et 50, et le gaz entre 50 et 120, ce qui explique la traditionnelle compartimentation verticale des ré-

servoires en zones saturée d'eau, de pétrole et de gaz de bas en haut. Si les trois fluides sont présents, il est préférable que le forage atteigne la couche de pétrole pour que celui-ci sorte avec une pression élevée, car cette couche est comprimée verticalement par le gaz et l'eau. Le pétrole sera économiquement plus viable s'il a un API élevé, car il sera plus facilement extractible, transportable et transformable dans une raffinerie. Pour comparaison, un pétrole d'excellente qualité a un API de 45.

2.3 Etat de l'art des études géophysiques effectuées sur Valhall

La première image sismique réalisée sur Valhall remonte à 1967, sur laquelle apparaissait au centre une "no data zone" liée au nuage de gaz, mais qui avait été initialement interprétée comme un dôme de sel. Cette interprétation fut corrigée lorsqu'une grille de données sismiques haute résolution fut acquise en 1975 et 1978. Des puits d'exploration confirmèrent le potentiel du réservoir en 1982, puis une autre campagne de sismique réflexion multitrace fut réalisée en 1983 pour enregistrer jusqu'à 3 km d'offset. On constata que le rapport signal sur bruit était meilleur pour les traces de faibles dépôts, et que les réflexions issues du réservoir ne suivaient pas un comportement hyperbolique. Le traitement des premiers 1.5 km d'offset améliora alors l'image de la zone centrale. Plusieurs acquisitions VSP furent ensuite testées avec succès, grâce à l'échantillonnage des réflecteurs par des trajets subverticaux à la fois descendants et montants. Il fut alors possible d'imager des réflecteurs profonds sous le nuage de gaz de manière continue. De plus, le gain en résolution permit de distinguer des failles, ainsi que le réflecteur correspondant au réservoir Hod, et de calibrer les profondeurs des structures imagées par la sismique (Munns (1985), figure 19 de l'article).

En 1990, une acquisition sismique 3D fut réalisée pour améliorer la couverture sismique des bords du champ, qui commençaient à être exploités suite à l'émergence des forages horizontaux (O. Barkved, 2003). Cette campagne confirma les observations de puits et démontra le fort potentiel de la zone nord du champ. En 1997 l'une des premières campagnes sismiques 3D fut réalisée à partir d'un réseau dense d'OBS 4 composantes Rosland *et al.* (1999). Le but était d'exploiter les ondes converties PS pour une meilleure imagerie de la zone centrale sous le nuage de gaz, les ondes S étant moins atténuées par le gaz que les ondes P. Une autre campagne 3D en 2002 montra l'impact de la réponse sismique 4D, sensible aux changements mineurs de contrainte dans le réservoir. L'analyse de l'AVOA² de ces jeux de données 3D par Hall *et al.* (2002); Hall & Kendall (2003) a mis en évidence l'anisotropie azimutale du sommet du réservoir. Cette anisotropie est directement reliée au réseau de fractures local, et représentative des contraintes tectoniques à grandes échelles qu'a subies la zone (Figure 2.5a). Les fractures et failles sont principalement orientées NO - SE, ce qui indique que les directions de contraintes à grandes échelles (d'extension puis de compression) étaient d'axe SO-NE.

Les données 3D-4C ont aussi été exploitées pour faire une étude de décalage des ondes S ('*shear wave splitting*') pour caractériser l'anisotropie azimutale (et donc les fractures) au niveau du plancher marin et au sommet du réservoir (Olofsson *et al.*, 2003). Cette méthode est basée sur le principe de biréfringence de l'onde S, qui en présence d'un milieu anisotrope se voit séparée en deux ondes S_V et S_H aux polarisations orthogonales et aux vitesses différentes. La figure 2.5b montre la corrélation entre l'anisotropie azimutale de la proche surface obtenue

2. Amplitude Versus Offset and Azimuth, basée sur l'équation de réflectivité fonction de l'angle d'incidence θ sur un réflecteur et de l'azimut ϕ : $R(\theta, \phi) = A + (B + D\cos 2\phi)\sin^2\theta + (C + E\cos 2\phi + F\cos 4\phi)\sin^2\theta \tan^2\theta$

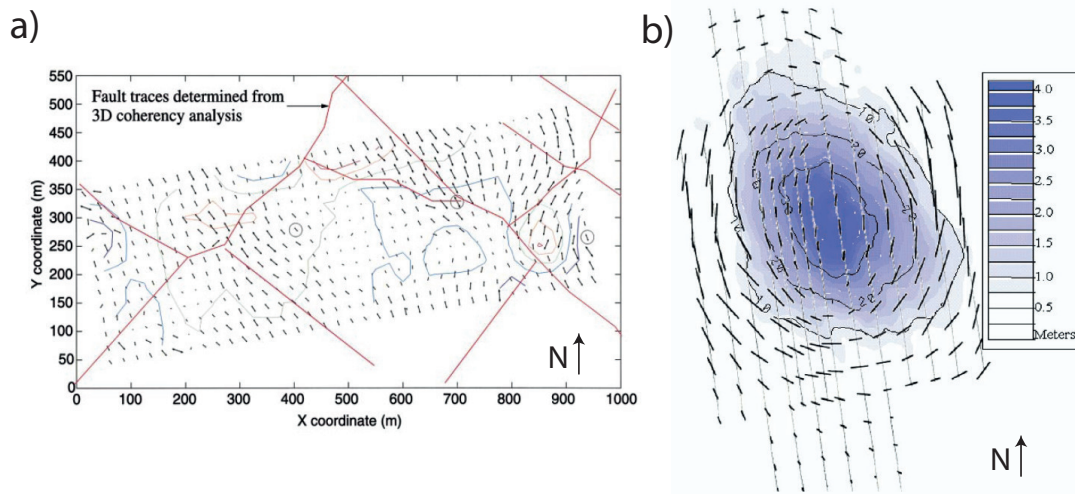


FIGURE 2.5 – (a) Correspondance au niveau du réservoir entre les failles (traits rouges) obtenues par analyse de cohérence 3D, et la carte du réseau de fractures, les traits noirs, dont la longueur est fonction du pourcentage d’anisotropie azimutale. Extrait de l’étude AVOA de Hall & Kendall (2003). La direction globale des failles et fractures est NO-SE. Les directions des fractures changent localement au niveau des failles (coin *est* par exemple). (b) Superposition de la carte de subsidence en très proche surface, et de la carte d’anisotropie azimutale (traits noirs) déduite de l’étude de Olofsson *et al.* (2003) sur les décalages d’ondes S.

par cette méthode, et la subsidence enregistrée dans la zone. L’anisotropie circulaire observée semble être due aux contraintes latérales exercées par la subsidence. La prise en compte de cette anisotropie azimutale améliora nettement l’image migrée obtenue à partir des ondes converties.

En 2003 fut mis en place un dispositif d’acquisition à 45 M \$ baptisé ‘LoFS’ (*Life of Field Seismic*), en raison de la répétition tous les six mois de l’acquisition sismique (Figure 2) (Kommedal *et al.*, 2004). Il est composé de 2302 récepteurs 4C disposés dans des câbles enterrés à 1 m de profondeur, et 50000 tirs environ sont effectués à chaque acquisition par canon à air (Figure 3.1a). La multiplication des jeux de données est utilisée pour des applications d’imagerie sismique 4D (van Gestel *et al.*, 2011; Huang *et al.*, 2011). Une discussion sur la sensibilité de l’imagerie 4D à l’acquisition, au prétraitement et à la manière de comparer de multiples jeux de données est présentée par Kommedal *et al.* (2007a); Kommedal & Barkved (2005) et Kommedal (2008). Cette technique s’est avérée être très sensible à la déplétion induite par la production, à la compaction du réservoir et à la subsidence de l’ensemble des sédiments sus-jacents au réservoir. Elle est très utilisée pour surveiller les variations de contraintes pouvant être responsables de l’effondrement des puits, et pour visualiser les migrations de fluides au sein du réservoir. Les migrations de fluides sont en partie provoquées par l’injection d’eau en profondeur afin d’optimiser la récupération de pétrole au niveau de puits de production.

La méthodologie pour déterminer les paramètres pétrophysiques du réservoir à partir de la sismique 4D est décrite sur la Figure 2.6 (van Gestel *et al.*, 2011, 2008a). Il est tout d’abord calculé des différentiels de temps d’arrivée et d’amplitude entre les différentes campagnes sismiques. Sur le champ de Valhall, les variations d’amplitudes sont reliées au durcissement du réservoir et à son amincissement, alors que les variations de temps de trajet sont dues aux

'W' de la Figure 2.6). En prenant d'autre part la relation $E_{ZZ} = \Delta z/z$, on obtient la relation : $\Delta t/t = (1 + R) E_{ZZ}$ (Hatchell *et al.*, 2005). L'intégration en z de cette formule de la surface au réservoir fournit une carte de compaction au niveau du réservoir. Les cartes de compaction et les cartes de différences d'amplitude observées sont ensuite comparées avec les cartes correspondantes calculées grâce à un code de modélisation du réservoir et de modélisations des ondes sismiques. La comparaison des cartes se fait automatiquement via un opérateur de comparaison (boîtes R et U, 'match quality factor'), et le critère de similarité des cartes permet d'évaluer la qualité d'un modèle de réservoir parmi les dizaines de modèles explorées.

Une extension de la méthode à l'utilisation des volumes de fluides extraits et injectés par les puits, a été employée par Huang *et al.* (2011); Barkved *et al.* (2009). Pour cela, la signature des différences entre les jeux de données ΔA_k pris à différents intervalles de temps ΔT_k , est reliée aux propriétés physiques G et aux volumes cumulatifs des différents types de fluides $\Delta V_j(\Delta T_k)$ injectés et extraits par les puits j dans la zone considérée : $\Delta A(x, y, \Delta T_k) = G \Delta V_j(\Delta T_k)$. Le terme G contient des informations sur la géologie, les propriétés des fluides, le modèle pétro-élastique, la connectivité et le degré de compartimentation. Le résultat est une carte de corrélation entre ΔA et ΔV permettant de délimiter les zones de drainages associées à un groupe de puit.

En 2004, une étude menée par Western Geco de tomographie anisotrope 3D en réflexion, a abouti à la création de macromodèles (3D) de vitesse d'ondes P (verticale), d'ondes S et des paramètres anisotropes de Thomsen δ et ϵ (Figure 3.2(a-c)). Cette méthode repose sur l'inversion conjointe des événements PP et PS, couplée à l'aplatissement des profils en images communes (*Common Image Gathers*, CIG's). L'utilisation de logs de puits pour calibrer la profondeur des réflecteurs est essentielle pour estimer les modèles anisotropes. Cette méthode de tomographie des temps nécessite une interprétation correcte dans les données des événements PP et PS. Le pointé des événements peut se faire dans les données avant sommation dans le domaine des temps ou des profondeurs, ce dernier domaine permettant parfois de pointer plus facilement (Stopin & Ehinger, 2001). Une démigration est alors faite pour obtenir les informations cinématiques du pointé profondeur, avec le même modèle de vitesse que celui utilisé pour la migration. En général, ce type de méthode essaye d'expliquer séquentiellement les données en considérant des arrivées de plus en plus tardives, afin de mettre à jour le modèle dans des couches de plus en plus profondes (*layer stripping approach*). Une application similaire à celle employée sur Valhall est décrite pour un autre cas d'étude de Mer du Nord par Broto *et al.* (2003). L'utilisation des modèles obtenus avec cette méthode comme modèles initiaux à mes applications d'inversion des formes d'ondes sur Valhall fut primordiale. J'illustre sur la Figure 2.7b la résolution et l'hétérogénéité du paramètre η (équation 1.11), qui est représentatif de l'anisotropie du milieu.

Le champ de Valhall est bien mieux contraint depuis 2009, date à laquelle une inversion des formes d'ondes en 3 dimensions a été réalisée par Sirgue *et al.* (2010). Nous montrons sur les Figures 2.8, 3.1b et 3.6f un extrait de ses résultats pour illustrer le détail avec lequel les structures sont imagées. Le modèle de vitesse haute résolution obtenu a permis de faire des interprétations stratigraphiques et géo-structurales du milieu, comme la détection de chenaux ou de fractures remplies de gaz. Son utilisation pour les applications d'imagerie 4D a drastiquement amélioré leur résolution. Des migrations de fluides peuvent être observées au sein du réservoir, et une meilleure planification du forage des puits peut être faite. Les modèles issus de l'inversion des formes d'ondes sont maintenant utilisés comme modèles initiaux aux techniques de migrations,

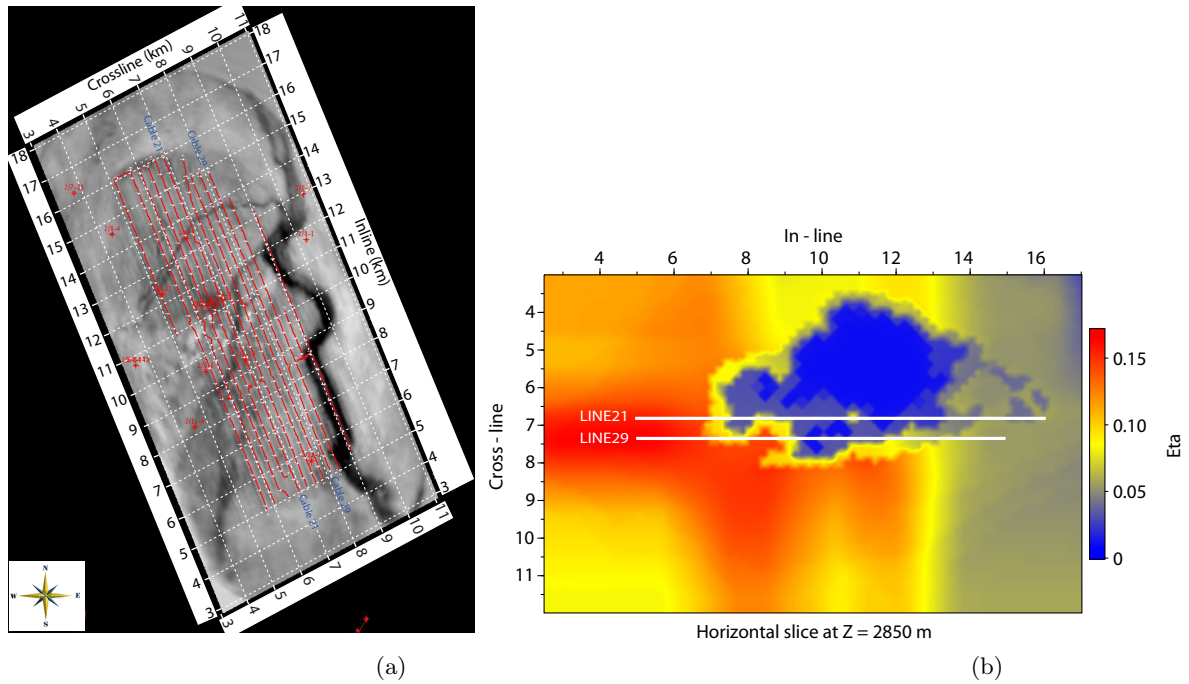


FIGURE 2.7 – (a) Carte indiquant la position des puits (étoiles rouges) et des récepteurs OBC (lignes rouges) dans le référentiel cartésien utilisé. Les OBC sont disposés le long de l'axe NO-SE de l'anticlinal. Le fond de cette carte correspond à la coupe horizontale extraite du modèle obtenu par FWI 3D à la profondeur de 150 m (Sirgue *et al.*, 2010), sur laquelle on reconnaît d'anciens chenaux (Copyright BP, used by permission). (b) Coupe horizontale du paramètre η obtenu par tomographie en réflexion anisotrope 3D sur le champ de Valhall, extraite à la profondeur de 2.85 km. La variabilité azimutale de ce paramètre n'est significative que dans la gamme de profondeur [2.2-3.3] km. Notez la forme bleue allongée, imageant le toit du réservoir dans lequel le paramètre η est nul.

comme illustré sur la Figure 2.8b, où l'on peut voir sur l'image migrée le réflecteur profond sous 3.5 km de profondeur, bien qu'il soit localisé sous le nuage de gaz. Bien que les réflecteurs et les composantes hautes fréquences de l'image soient atténuées dans le nuage de nuage et en dessous, les réflecteurs sont bien plus marqués que si un modèle de tomographie standard avait été utilisé (Sirgue *et al.*, 2010).

Pour avoir un aperçu général de la diversité des méthodes employées sur Valhall, on peut aussi citer des études réalisées par interférométrie de bruit sismique enregistré par le réseau LoFS (Stewart, 2006; Artman, 2007; Dellinger, 2008; Dellinger & Yu, 2009). Le succès de ces applications a résidé dans l'utilisation de données dans une gamme appropriée de fréquences [0.1;2] Hz, pour laquelle le bruit sismique est omnidirectionnel³. Le bruit en dessous de 0.1 Hz est issu de la houle, il est très présent sur l'hydrophone et quasi-absent sur la composante verticale de vitesse. L'énergie microsismique aux fréquences d'intérêt est générée par l'interaction des vagues et du plancher marin, ce qui explique pourquoi ce bruit est de *meilleure qualité*

3. condition nécessaire à l'application de la technique d'interférométrie

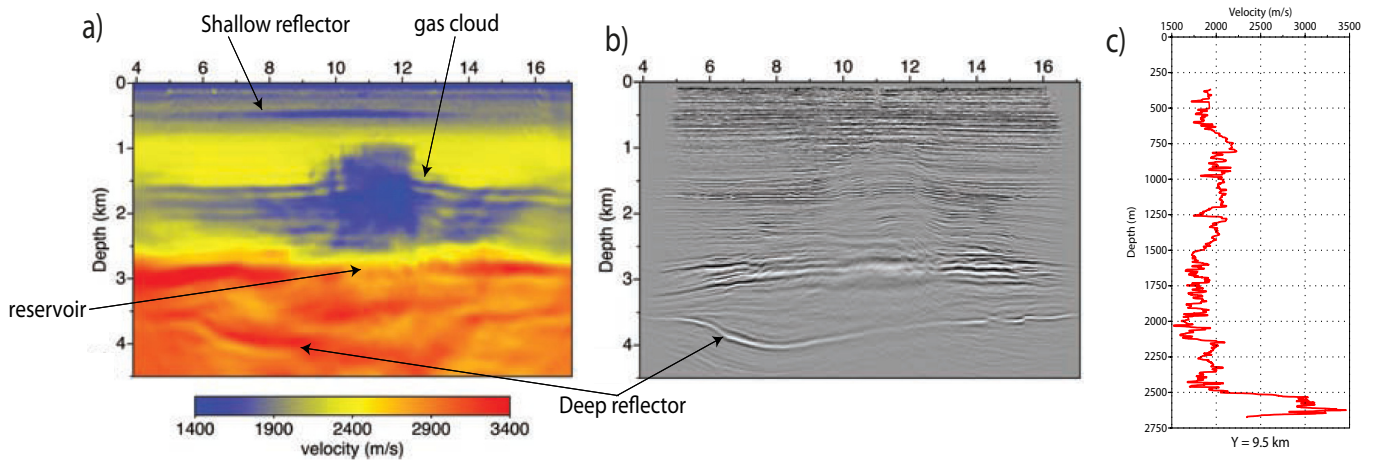


FIGURE 2.8 – (a) Coupe verticale extraite du modèle obtenu par inversion des formes d’ondes à trois dimensions par Sirgue *et al.* (2010) à la position $X = 5.575$ m. (b) Image obtenue par migration avant sommation utilisant l’équation d’onde ‘one-way’, calculée dans le modèle montré en (a). (c) Log sonique extrait au niveau de la ligne 21 à la position ($X = 6.8$ km ; $Y = 9.5$ km), représentatif d’une vitesse verticale.

lorsqu’il est enregistré par temps orageux. Le bruit de fréquence supérieure à 2 Hz est issu des activités humaines de la plateforme, généré par le passage de bateaux ou des acquisitions sismiques distantes. Par exemple, Ridder & Dellinger (2011); Mordret *et al.* (2011) ont pu obtenir par tomographie de bruit sismique une image de la très proche surface (< 200 m) des vitesses d’ondes P et d’ondes de Scholte (voir partie 6.9.1). D’autre part, Chambers & Kendall (2010) ont pu localiser dans l’espace les événements sismiques induits par de la fracturation hydraulique.

2.4 Présentation du jeu de données OBC

2.4.1 Anatomie des données

Dans un milieu à deux couches, avec un gradient de vitesse dans chacune des couches, on distingue trois principaux types d’onde : l’onde directe, l’onde réfléchie et l’onde réfractée. Par la suite, je fonderai mon interprétation qualitative des données Valhall sur ce type d’arrivée. La Figure 2.9a illustre comment se propagent ces trois types d’ondes. Deux instantanés des fronts d’ondes à environ 2 s et 5 s après l’émission de la source sont montrés. Au moment du deuxième instantané apparaît la première onde réfractée, apparaissant au niveau de la distance critique. Celle-ci est fonction de l’angle d’incidence critique de l’onde, dépendant du contraste de vitesse entre les deux couches, et de la profondeur de l’interface. Les ondes transmises dans le milieu inférieur sont transmises une deuxième fois dans le milieu supérieur lors de leur trajet ascendant (d’où le nom d’ondes tournantes), grâce à l’augmentation de vitesse avec la profondeur dans les couches. Les temps d’arrivée des ondes réfractées et réfléchies sont représentés dans le cas de deux couches homogènes sur un graphique ‘temps-offset’ (Figure 2.9b). Ces courbes sont appelées hodochrones. La première couche a une vitesse V_0 et une épaisseur h_0 , la deuxième

couche une vitesse V_1 . X_C est l'offset critique à partir duquel les premières ondes réfractées apparaissent, et le point de brisure est la distance à partir de laquelle l'onde conique arrive avant l'onde directe. On voit que la pente constante de l'hodochrone de l'onde directe renseigne sur la vitesse V_0 de la première couche. Le temps d'arrivée de l'onde réfléchi pour l'offset nul est $t = (2h_0)/V_0$. Cela illustre de manière schématique l'indétermination à court offset entre l'épaisseur de la couche et sa vitesse pour expliquer le temps d'arrivée. À grand offset, l'hodochrone de la réfléchi tend vers celle de l'onde directe, d'équation $t = x/V_0$, et porte donc l'information sur la vitesse. C'est pour cela que l'approche hiérarchique classique consiste à inverser d'abord les grands offsets pour déterminer la vitesse, pour ensuite éviter l'ambiguïté vitesse-profondeur lors de l'inversion des courts offsets.

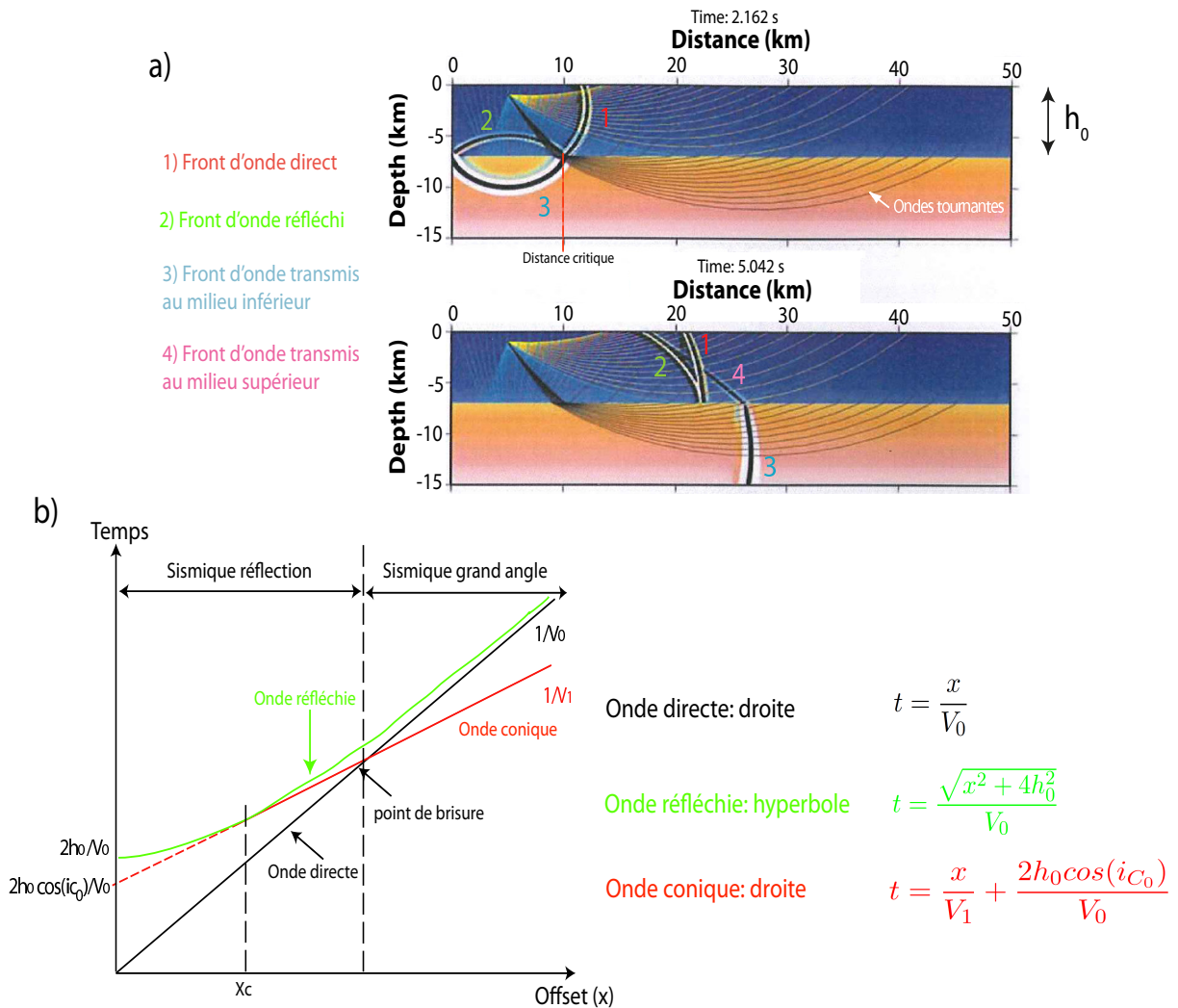


FIGURE 2.9 – (a) Deux instantanés de propagation dans un milieu bi-couches. (b) Hodochrones des trois principales arrivées : onde directe, réfléchi et réfractée. L'onde réfléchi tend asymptotiquement vers l'onde directe à grand offset.

2.4.2 Composante de pression calculée pour le cas synthétique de Valhall

Nous présentons maintenant la signature des structures géologiques du modèle synthétique de Valhall (Figure 2.2a) dans les données sismiques enregistrées en pression par les hydrophones (Figure 2.2d). On reconnaît sur le profil en tir commun l'onde directe à sa pente constante égale à $(1/1500)$ s/m. On observe aussi sur le profil une démultiplication du front d'onde, phénomène connu sous le nom de "shingling" (Robertson *et al.*, 1996), se traduisant par l'apparition de nouveaux cycles et un pic d'énergie qui se déplace ultérieurement dans le front d'onde. L'origine de ces ondes dispersives serait ici la présence d'un gradient de vitesse, qui agirait comme un guide d'onde pour les arrivées réfractées (Sheriff & Geldart, 1995).

Par ailleurs, on constate qu'à chaque interface I_i correspond une réfléchie R_i et une réfractée D_i . Les rais calculés dans le modèle synthétique permettent de voir que les ondes réfractées se propagent principalement le long de 4 interfaces dont la profondeur est de haut en bas : 0.2, 0.7, 1.1 et 2.7 km de profondeur. Les premières arrivées sont constituées de réfractées de plus en plus profondes au fur et à mesure que l'offset augmente. La position X_{i-j} indique la distance à laquelle l'onde se réfractant le long de l'interface j devient plus rapide que celle se réfractant sur l'interface i . Le léger décalage de la marque X_{3-4} sur le modèle et le profil vient du fait que les rais pour le premier ont été calculés dans le modèle exact, et que les temps de premières arrivées ont été calculés pour le second dans une version lissée du vrai modèle, en raison de la précision limitée du tracé de rais dans les modèles trop contrastés. Aucune première arrivée ne se propage dans le nuage de gaz à cause des basses vitesses le caractérisant. Le calcul de sismogrammes avec plus de déport a montré qu'il fallait 24 km d'offset pour que des ondes se propageant sur la cinquième interface à 5 km de profondeur soient enregistrées en première arrivée. Cette distance est importante en raison de la couche plus basse vitesse séparant les deux interfaces de haute vitesse à 2.7 km et 5 km de profondeur.

2.4.3 Données 4C du dispositif LoFS

Avant de présenter notre interprétation des phases observées sur les données 4 composantes, nous repositionnons sur la Figure 2.7a le référentiel cartésien utilisé dans les en-têtes sismiques par rapport à l'orientation de l'anticlinal d'axe NO-SE. Le dispositif d'acquisition comprend un offset maximum d'environ 13 km, l'espacement entre sources et entre récepteurs est de 50 m, et la distance séparant deux câbles est de 300 m. Le temps d'enregistrement est pour tous les offsets de 8 s à partir du moment où le canon émet la source. Notons que les jeux de données enregistrés sur les champs pétroliers sont parfois pollués par des acquisitions sismiques voisines. Des techniques pour atténuer ces interférences, plus efficaces que l'application d'un simple filtre en k , tirent parti de la répétabilité des intervalles de tirs pour les deux acquisitions en conflits, pour soustraire au signal utile le signal parasite enregistré après la fenêtre de temps utile (Kommedal *et al.*, 2007b).

Le prétraitement de la composante d'hydrophone a consisté en l'omission manuelle des traces bruitées, en l'application d'un filtre de pente pour enlever les ondes d'interfaces (visibles dans le premier kilomètre d'offset, Figure 2.10a) ; puis d'une déconvolution spectrale (blanchiment) afin de compresser la forme d'onde et de résoudre les réflexions faiblement espacées. Je montre l'effet de ce dernier prétraitement dans le domaine des temps et des fréquences sur les Figures 2.10(a-b) et 2.11(a-b).

L'aplatissement du spectre permet aussi de considérer des groupes de fréquence dont les composantes ont des amplitudes voisines lors de l'inversion des formes d'ondes. On observe sur

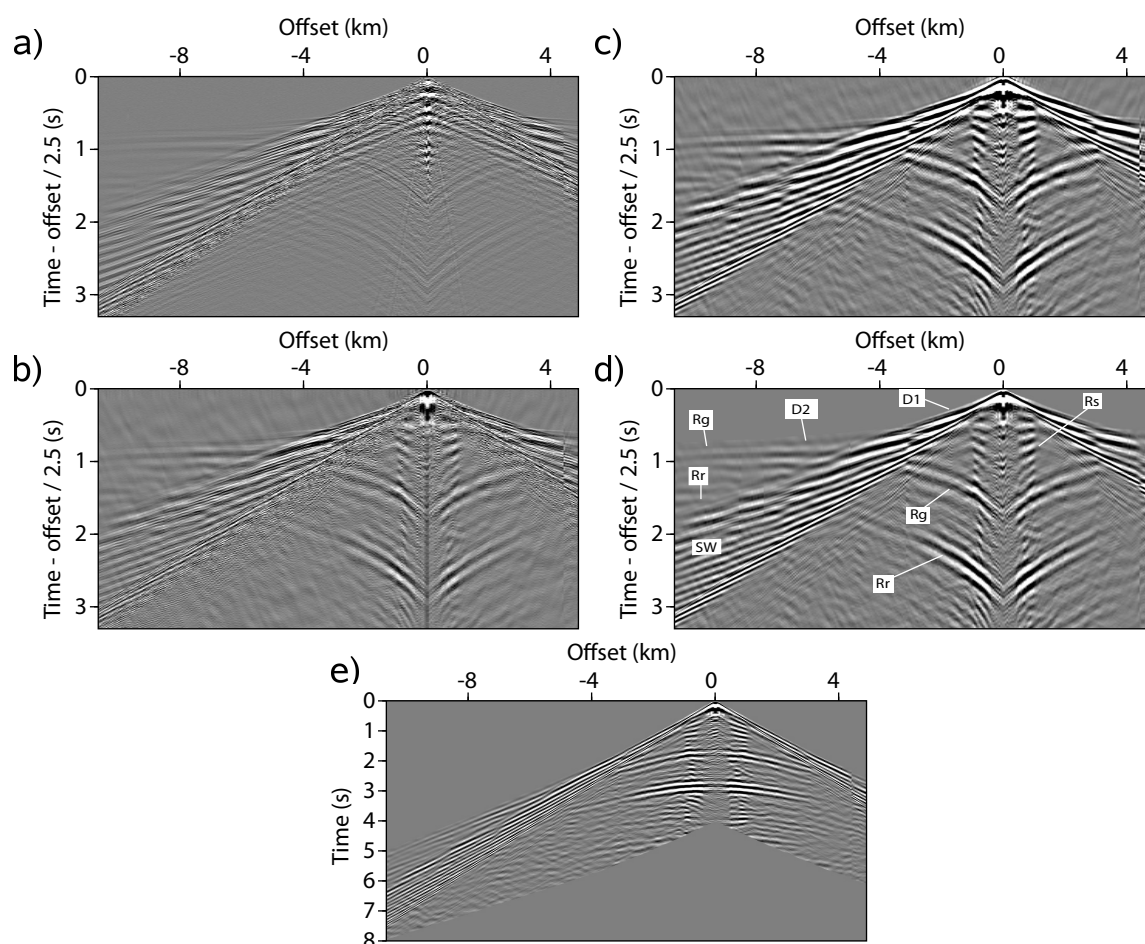


FIGURE 2.10 – Profil en tir commun de la ligne 29 à différents stades du prétraitement. Dans l'ordre du prétraitement, nous présentons les données brutes (a) ; et après application (b) d'un filtre de pente et d'une déconvolution, (c) d'un filtrage butterworth entre 3 et 20 Hz, et (d) d'un filtre de cohérence par matrice spectrale, d'une loi de mute et d'une correction d'amplitude 3D - 2D suivant une loi en \sqrt{t} (produit final). En (d) est indiqué le nom des principales arrivées : D1, D2 = ondes réfractées (*diving waves*) à une profondeur inférieure à 0.7 et 1.25 km respectivement ; Rg, Rr = ondes réfléchies au toit du gaz et au toit du réservoir ; SW = train d'onde dit *shingling*. (e) Même profil qu'en (d), mais représenté sans vitesse de réduction.

le profil en récepteur commun de la Figure 2.10b que les composantes basses fréquences du signal sont plus visibles après déconvolution. Un filtre passe-bande Butterworth est ensuite appliqué entre 3 et 20 Hz pour obtenir un meilleur rapport signal sur bruit, le blanchiment ayant aussi augmenté la part de bruit haute fréquence. On voit que l'effet du filtre, Figure 2.10c, donne un meilleur rendu visuel grâce à l'omission du bruit haute fréquence, bien que celui-ci n'ait pas d'incidence sur l'inversion qui ne considère que des fréquences plus basses. Un avantage de ce filtre est qu'il ne crée pas de précurseurs. Après cela, l'utilisation d'un filtre de cohérence par matrice des intercorrélations dans le domaine des fréquences (matrice spectrale) (Glangeaud & Coppens, 1997) permet d'améliorer la cohérence latérale des événements dont l'offset est supérieur à 4 km. Il n'est pas appliqué aux offsets inférieur car le filtre tendrait à détruire les

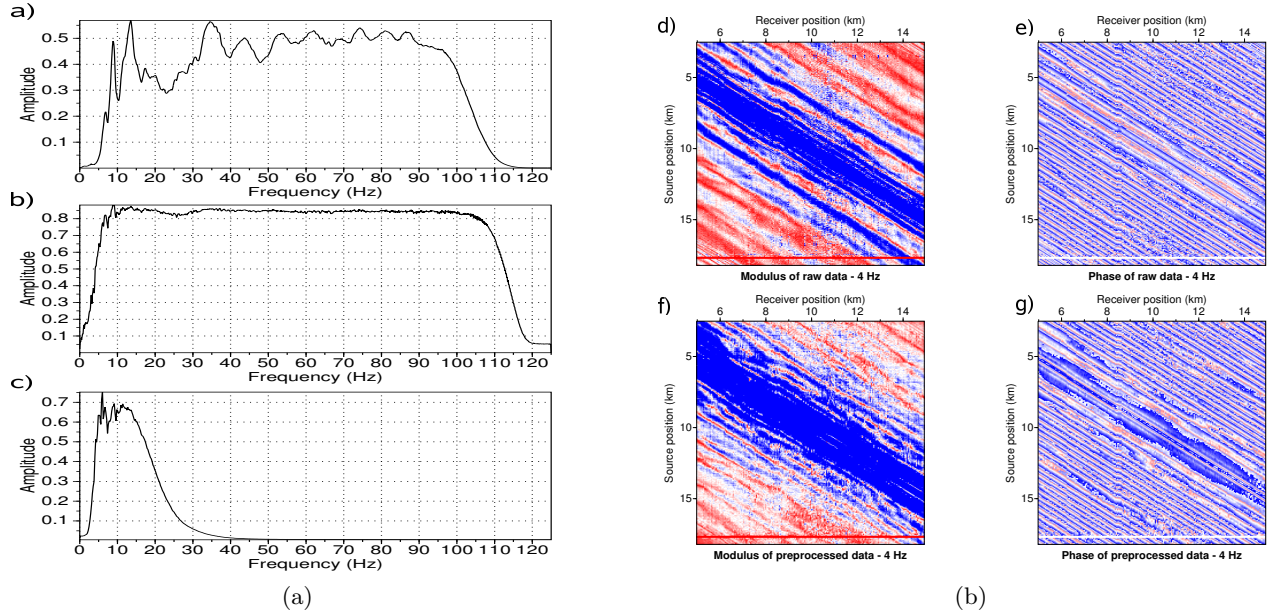


FIGURE 2.11 – (a-c) Spectre fréquentiel des profils en récepteur commun présentés sur la Figure 2.10(a,b,d) respectivement. Notez l’aplatissement du spectre après déconvolution (b), et l’effet du filtre passe-bande appliqué aux données (c). (d-g) Carte du module (d,f) et de la phase (e,g) à 4 Hz du jeu de données de la ligne 29, avant (d-e) et après (f-g) l’ensemble du prétraitement décrit dans le texte.

réflexions présentes à ces dépôts. La matrice spectrale est composée de tous les autospectres et interspectres possibles entre les diverses traces. La projection des données sur les premiers vecteurs propres de la matrice spectrale fournit l’espace signal. Ce traitement est généralement utilisé comme technique de séparation d’onde. Ici, en utilisant des moyennes en fréquence et en distance faibles, nous l’utilisons comme outil pour améliorer le rapport signal sur bruit (Ravaut *et al.*, 2004). Précisons qu’un AGC est appliqué avant ce traitement pour optimiser l’utilisation du filtre, et que les vraies amplitudes sont restituées a posteriori. Par ailleurs, un mute est appliqué avant la première arrivée et après les 4 s suivant la première arrivée, ce dernier mute permettant de ne pas prendre en compte les arrivées tardives plus complexes. Les données sont aussi multipliées par une fonction en \sqrt{t} pour imiter la décroissance géométrique des ondes modélisées par une approche 2D. On notera que l’utilisation de cette loi peut potentiellement fausser le comportement de l’AVO élastique des données. Il existe des filtres latéraux plus appropriés, notamment développés par Wapenaar *et al.* (1992).

Les profils de la Figure 2.10(a-d) sont représentés avec une vitesse de réduction de 2500 m/s. L’axe des temps (t') d’un profil avec une vitesse de réduction est une fonction de l’offset (x) et du temps (t) suivant la loi : $t' = t - x/v_{red}$, où v_{red} est la vitesse de réduction. Je fais la comparaison des deux représentations sur la Figure 2.10d,e. Le profil représenté avec vitesse de réduction (Figure 2.10d) permet de conclure que l’onde réfractée D2 a une vitesse moyenne de 2500 m/s en raison de sa pente horizontale constante.

J’illustre sur la Figure 2.11(d-g) l’effet de la chaîne de prétraitement sur les cartes du module et de la phase du jeu de données à 4 Hz. On observe des différences notables entre les cartes de

phase au niveau de la bande diagonale pour le premier kilomètre d'offset, qui pourrait être dû à l'omission des ondes de surface par le filtre de pente. De plus, l'espacement entre les franges de cohérence est plus grand dans la carte de phase après prétraitement, en particulier pour les grands offsets.

Il nous faut, pour faire l'interprétation de la composante d'hydrophone, une idée générale de la structure du milieu. Nous faisons référence pour cela au modèle initial de vitesse construit par tomographie en réflexion (Figure 3.2a), ainsi qu'au log sonique de vitesse présenté sur la Figure 2.8c. Comme pour le modèle synthétique, le milieu réel est relativement tabulaire, il comporte une zone à faible vitesse associée à une accumulation de gaz en subsurface et une interface haute vitesse à environ 2.6 km de profondeur. Les principales arrivées visibles sur l'hydrophone sont les ondes réfléchies au toit du gaz et du réservoir, à 1.5 km et 2.6 km de profondeur, indiquées sur les données d'hydrophone de la Figure 2.10d par les lettres "Rg" et "Rr" respectivement. Les premières arrivées ne se propagent que dans les premiers 1.25 km de profondeur. Le trajet emprunté par ces différentes ondes est illustré sur la Figure 3.3b. L'onde réfractée sur le toit du réservoir est enregistrée comme première arrivée à des offsets supérieurs à 14 km (non disponibles), contre 11 km d'offset pour le cas synthétique.

Il existe sur toutes les composantes du jeu de données un bruit directionnel associé à la présence de la plateforme, dont l'intensité est fonction de la distance à celle-ci (Figure 2.12a). La comparaison de deux profils en récepteur commun sur la ligne 21 (Figure 2.12(b-c)) illustre à quel point ce bruit peut *masquer* l'information utile. On s'aperçoit que ce bruit laisse une empreinte hyperbolique dans les profils triés en tirs communs, laquelle hyperbole est d'autant plus étalée en espace et faible en amplitude que l'on s'éloigne de la plateforme (comparaison des profils de la ligne 21 et 29 Figure 2.12(d-e)). Pour les récepteurs communs proches de la plateforme, le bruit apparaît de manière identique pour toutes les traces car le signal est, quelle que soit son origine, noyé sous le bruit. Par contre sur les profils en tirs communs, seules les traces correspondant aux récepteurs proches de la plateforme sont bruitées.

Nous présentons sur la Figure 2.13e l'orientation des composantes V_X et V_Z des données de géophones par rapport au référentiel cartésien utilisé, lequel est défini par une direction *In-line* et *Cross-line* (respectivement parallèle et perpendiculaire aux lignes OBC). Ces composantes ont pour vocation d'enregistrer les ondes de compression et les ondes SV, exploitées par notre code d'inversion des formes d'ondes élastique.

Le prétraitement des données de géophone comporte quelques différences par rapport à celui utilisé pour l'hydrophone. La déconvolution spectrale, le filtrage butterworth, le mute avant la première arrivée et la multiplication des données par \sqrt{t} sont toujours appliqués. Par contre le filtre de pente et le filtre de cohérence par matrice spectrale sont abandonnés car nous voulons exploiter les ondes S lentes à courts et moyens offsets, indiquées sur la Figure 2.13b par les hodochrones rouges (ondes converties SS) et la plus lente des hodochrones bleues (indiquant l'onde S réfractée dans la couche à [700-1500] m de profondeur).

L'interprétation des ondes converties dans les données est primordiale pour justifier la véracité des structures imagées par inversion des formes d'ondes, car celles-ci porteront une empreinte dans les données calculées qui doit être compatible avec les ondes observées. Une interprétation des ondes converties pour une précédente acquisition fut présentée par Kommedal *et al.* (1997) et Thomsen *et al.* (1997).

Pour savoir comment s'expriment dans les données les différentes ondes converties PP, SS et PS, nous avons utilisé le code de Podvin & Lecomte (1991) résolvant l'équation Eikonale, afin de trouver le trajet au temps de parcours minimum entre la source, l'interface et le récepteur. Pour

cela, nous avons pointé quatre interfaces principales dans les modèles de V_P et de V_S obtenus par inversion des formes d'ondes (chapitre 6), aux profondeurs de 0.7, 1.25, 1.5 et 2.6 km. Pour une onde convertie PS par exemple, il faut calculer la carte des temps de trajet dans le modèle V_P entre la source et tous les points de l'interface ; puis la carte dans le modèle V_S entre un récepteur et cette même interface. Le parcours de l'onde PS est trouvé en essayant toutes les possibilités de trajet ascendant et descendant, de sorte que le temps de trajet soit minimum (principe de Fermat). En raison de l'anisotropie VTI significative affectant le milieu, nous avons créé un milieu effectif isotrope pour le modèle de V_P afin qu'une modélisation isotrope puisse prédire relativement bien la cinématique des réflexions et des réfractions. Pour cela, nous avons transformé le modèle vertical de vitesse trouvé par inversion des formes d'ondes anisotrope en modèle horizontal. Ce modèle fut ensuite converti dans le domaine des temps, puis repassé dans le domaine des profondeurs mais en utilisant le modèle horizontal dont la vitesse a été augmentée de 5%. L'utilisation de ce modèle plus rapide permet de faire un étirement vertical de la structure, afin que les réflexions, qui étaient dans le modèle horizontal en avance, arrivent plus tard. L'utilisation du modèle horizontal de vitesse comme point de départ de la conversion profondeur-temps-profondeur est justifiée par le fait que celui-ci permet de prédire les temps de premières arrivées. Notons que la cinématique de ce modèle est peu sensible à *grands déports* aux étirements verticaux de structure. Pour le calcul des hodochrones, les interfaces seront donc déplacées vers le bas dans le modèle de V_P étiré, alors que nous considérons un modèle de V_S non étiré pour la branche ascendante, V_S n'étant pas sensible à l'anisotropie VTI en deux dimensions.

Les hodochrones des premières arrivées P et S, et des ondes converties sur les 4 interfaces mentionnées sont représentées sur la Figure 2.13b. Il est intéressant de voir que les hodochrones des premières arrivées des ondes P et S (en bleu) comportent toutes deux une brisure, à respectivement 3.5 et 2 km d'offset, bien que celle-ci soit difficile à observer sur la composante V_X pour l'onde S. Cette rupture de pente correspond à la distance à partir de laquelle la réfractée se propageant dans la couche à [700-1500] m de profondeur est enregistrée comme première arrivée. Aux offsets supérieurs à cette brisure, les arrivées enregistrées après cette réfractée et avant la réfractée des 700 premiers mètres de profondeur, correspondent aux ondes dispersives précédemment nommées '*shingling*'. Par ailleurs, nous observons des ondes lentes sur toutes les composantes, à l'intérieur du cône délimité par la plus lente des hodochrones bleues. Les ondes de cisaillement ne pouvant naturellement pas être enregistrées sur l'hydrophone, ces ondes sont interprétées comme étant des ondes d'interfaces fluide-solide de Scholte, dont la vitesse est inférieure à celle de l'onde S. Ce type d'onde est souvent enregistré dans les environnements marins, caractérisés par des sédiments mous au niveau du fond marin (Carcione & Helle, 2004).

Dans la nomenclature utilisée par la suite, j'omets la première lettre P des ondes converties PPP, PPS, PSS et PSP, correspondant à la propagation en mode P dans la couche d'eau. L'indice i des converties PP_i (par exemple) indique la profondeur (en km) à laquelle a lieu la conversion PP. Nous notons une correspondance entre les hodochrones des converties PS (en vert) et des arrivées latéralement cohérentes visibles sur le profil de la Figure 2.13b. Leur vitesse moyenne apparente est de 2500 m/s, car les hodochrones sont horizontales sur le profil. La conversion PS ayant la plus forte amplitude dans les données est celle se produisant au toit du réservoir à 2.6 km de profondeur. Nous ne voyons pas malheureusement pas d'arrivées latéralement cohérentes dans les données qui pourraient correspondre aux ondes SS (quelle que soit l'interface). Les données ne comportant que 8 s d'enregistrement à partir de l'émission de la source, la gamme de signal utile enregistrée se réduit avec l'offset. Nous n'avons pas la

possibilité d'observer de signal latéralement cohérent dans les données qui pourrait correspondre à la convertie $SS_{2.6}$, son hodochrone arrivant à 7.8 s à offset nul. Malgré un enregistrement de 11.5 s, Kommedal *et al.* (1997) n'a pas non plus été en mesure d'observer cette convertie dans son jeu de données.

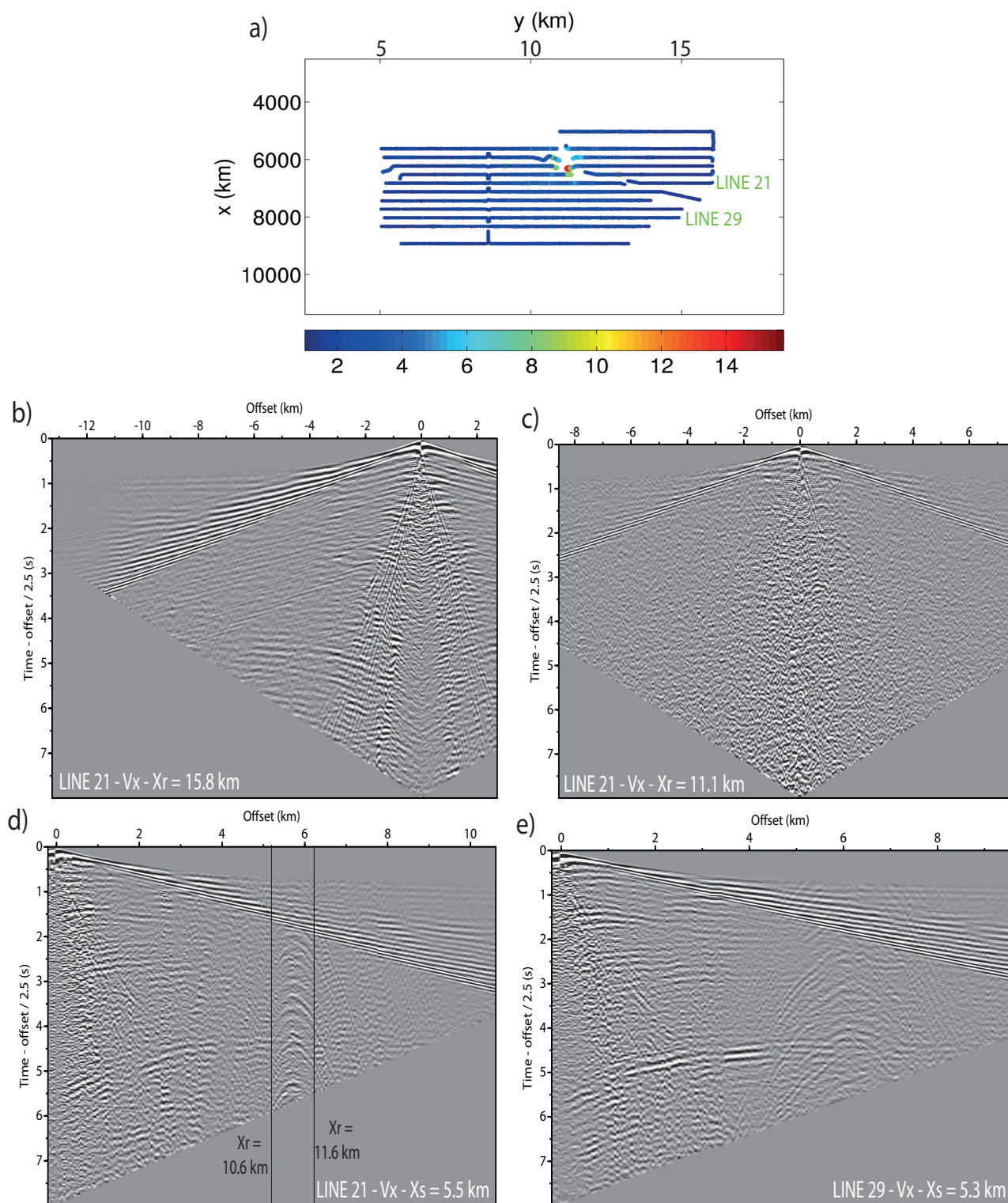


FIGURE 2.12 – (a) Carte indiquant le niveau de bruit pour chaque récepteur, établi comme le rapport entre la partie réelle (sommée sur les tirs) des données observées sur celle des données calculées par une modélisation 3D (courtoisie de Vincent Etienne et Hu Guanghui). Le bruit est localisé au centre de l'acquisition au niveau de la plateforme. (b-c) Profils en récepteurs communs extraits de la ligne 21 pour la composante V_X , situés sur le bord et le centre de l'acquisition respectivement. (d-e) Profils en tirs communs pour la ligne 21 et 29.

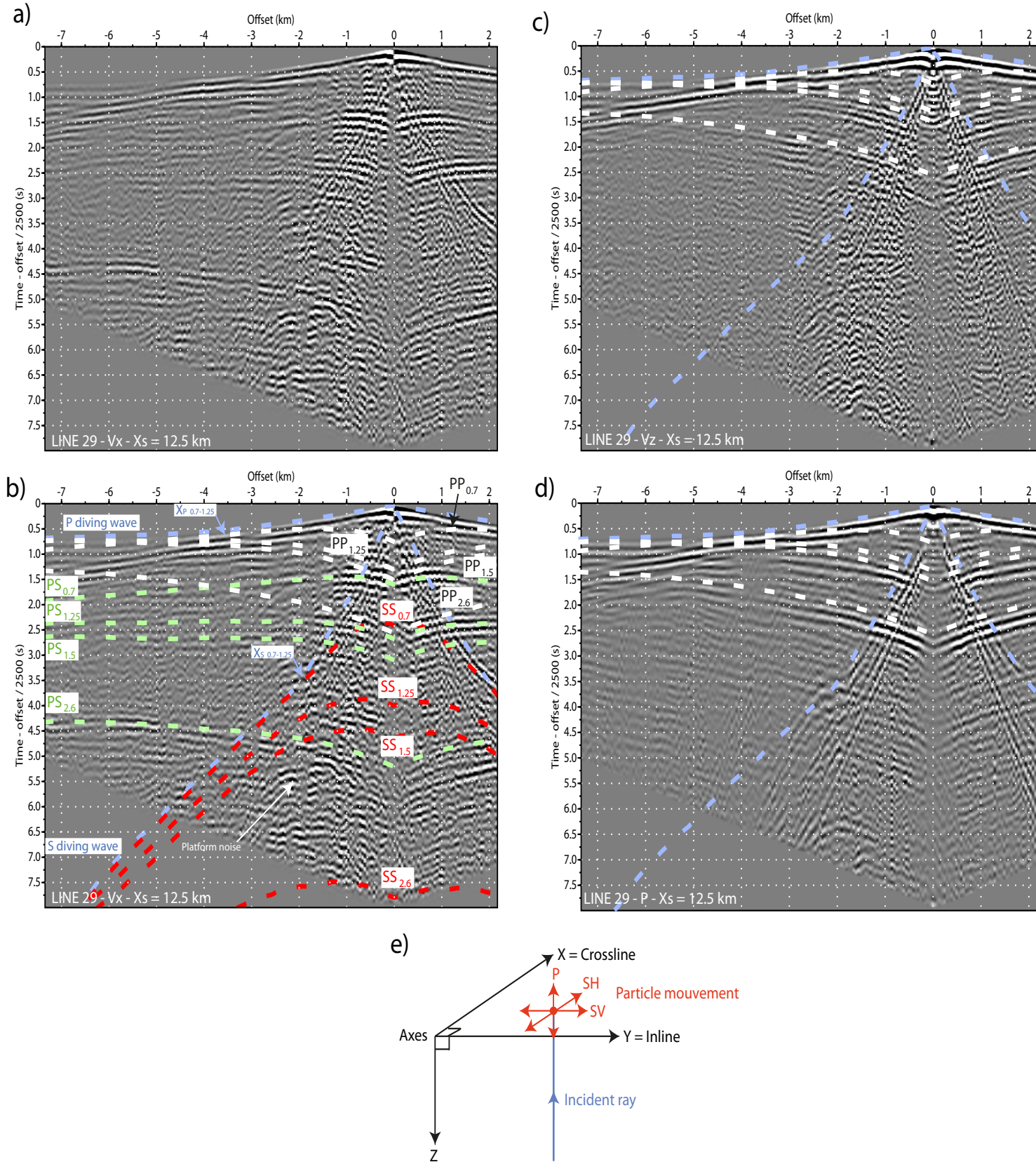


FIGURE 2.13 – Profils en tir commun issus de la ligne 29 pour les composantes V_X (a-b), V_Z (c) et de pression (d), filtrés entre 1 et 8 Hz. (b-d) Nous avons superposé en pointillé couleurs les hodochrones des ondes converties PP (en blanc), PS (en vert) et SS (en rouge), calculées par tracé de rais dans les modèles de V_P et V_S obtenus par inversion des formes d'ondes élastique (chapitre 6). (e) Schéma définissant le mouvement particulaire d'une onde dans notre système d'axe.

Chapitre 3

Inversion des formes d'ondes acoustique monoparamètre

Sommaire

3.1	Introduction au chapitre	89
3.2	Empreinte de l'anisotropie du milieu sur une inversion acoustique isotrope	90
3.2.1	Présentation de l'article	90
3.2.2	Résumé	91
3.2.3	Introduction	92
3.2.4	Méthodes	95
3.2.5	Application au champ de Valhall	97
3.2.6	Résultats complémentaires à l'article	121
3.2.7	Conclusion de l'article	124
3.3	Analyse paramétrique de l'inversion acoustique	126
3.3.1	Non-linéarités introduites par la surface libre - apport du précondition- nement - sensibilité à l'offset	126
3.3.2	Sensibilité aux modèles initiaux anisotropes, à la fréquence initiale et au préconditionnement des données, relativement à une inversion isotrope	127
3.3.3	Sensibilité de l'inversion aux imprécisions de la partie superficielle du modèle initial	133
3.3.4	Sensibilité à la méthode d'optimisation et à la régularisation	133
3.3.5	Sensibilité aux normes sur données réelles	136
3.4	Conclusion du chapitre	137

3.1 Introduction au chapitre

Ce chapitre est consacré à l'analyse de sensibilité de l'inversion acoustique à différents paramètres de réglage et approximations, étape nécessaire avant d'envisager des inversions multiparamètres. Dans la première partie de ce chapitre, je présente les résultats de l'inversion acoustique monoparamètre des données réelles enregistrées par les hydrophones du câble 21 de la campagne d'OBC réalisée sur le champ de Valhall (voir Chapitre 2). Ces résultats ont

été publiés en décembre 2011 dans un article du *Geophysical Journal International*. Dans cet article, je montre l’empreinte de l’anisotropie sur les résultats d’une inversion isotrope par comparaison avec les résultats d’une inversion anisotrope monoparamètre. Je complète les résultats de l’article par une analyse du préconditionnement de l’inversion par les termes diagonaux de deux approximations du Hessien, avant de résumer les enseignements tirés de cette étude sur le champ de Valhall.

Dans la deuxième partie du chapitre, je discute l’influence de plusieurs facteurs sur la qualité des résultats. Cette deuxième partie peut être vue comme une annexe à l’article, servant à justifier les stratégies adoptées lors de l’inversion des données réelles.

J’illustre tout d’abord, à l’aide du modèle synthétique de Valhall, les effets de non-linéarité à grands offsets induits par les multiples sur la surface libre. Je montre comment l’approche hiérarchique fondée sur l’utilisation de facteurs d’amortissement décroissants permet de réduire ces effets non-linéaires. Je m’intéresse ensuite au cas d’une inversion acoustique anisotrope où les modèles des paramètres de Thomsen sont des représentations lissées des vrais modèles et sont maintenus fixes lors de l’inversion. Je montre que le lissage des paramètres de Thomsen induit des erreurs cinématiques et dynamiques à grands angles. Ces erreurs provoquent l’apparition d’artefacts lorsque la fréquence initiale est de 3.5 Hz (comme dans le cas réel), qui peuvent être supprimés en utilisant un modèle de vitesse initial plus précis dans la partie superficielle. J’illustre ensuite sur le cas d’étude réel le rôle de la régularisation pour contrôler le rapport signal sur bruit du modèle reconstruit, lorsque la méthode d’optimisation de type L-BFGS est utilisée. L’influence des normes L1 et L2 est ensuite discutée à partir des résultats obtenus sur les lignes 21 et 29.

3.2 Empreinte de l’anisotropie du milieu sur une inversion acoustique isotrope

3.2.1 Présentation de l’article

On distingue dans la nature l’anisotropie intrinsèque, ou de "texture", produite par l’orientation préférentielle des minéraux, de l’anisotropie extrinsèque associée aux hétérogénéités du milieu. Cette dernière peut être liée à la fissuration mise en place dans un champ de contrainte anisotrope, lorsque les contraintes hydrostatiques sont inférieures à la pression de fermeture des fissures. L’anisotropie extrinsèque peut aussi être générée par un litage des sédiments dont l’épaisseur caractéristique est trop faible pour être résolue par une technique d’imagerie (Bakus (1962), Pratt *et al.* (2001)). Il est difficile de faire la distinction entre les deux types d’anisotropie, notamment car la notion d’hétérogénéité et d’anisotropie dépend de la fréquence considérée (Tsvankin, 2001). Le milieu de Valhall est caractérisé par une forte anisotropie VTI, comportant localement des vitesses verticales 15 % plus lentes que les vitesses horizontales. Les deux causes les plus probables de l’anisotropie VTI sont l’alignement préférentiel des minéraux dans les argiles (Sayers, 1994), et l’agencement périodique de fines couches isotropes aux propriétés différentes (Tsvankin, 2001). L’anisotropie observée dans les données sismiques est une anisotropie effective pouvant être associée à ces deux causes. Initialement, nous avons pu nous rendre compte de cette anisotropie en calculant des sismogrammes dans le modèle initial de vitesse fourni par BP, construit par tomographie en réflexion anisotrope et transformé en vitesse NMO (*Normal Move Out*) pour l’application de l’inversion des formes d’ondes isotrope.

Ce modèle de vitesse NMO permet de redresser les hyperboles de réflexion visibles dans les profils en points milieux communs. Les sismogrammes calculés dans le modèle NMO prédisent correctement les réflexions, mais nous avons noté un écart significatif entre les temps de trajet des premières arrivées observées et calculées sur les traces à grands déports. L'inversion des formes d'ondes étant basée sur l'utilisation conjointe des données enregistrées à faibles et à grandes ouvertures, il nous a paru intéressant de créer un modèle initial fournissant un meilleur ajustement des temps de trajet des premières arrivées. Nous avons pour cela réalisé une tomographie des temps de première arrivée à partir d'un pointé manuel de ces phases (effectué lors de mon stage de Master 2), aboutissant au modèle dit FATT¹. Pour cela, le modèle NMO a été utilisé comme modèle initial car sa résolution est supérieure à ce que nous pouvions espérer d'une tomographie de première arrivée partant d'un modèle à gradient de vitesse. Le modèle initial FATT nous paraissait potentiellement plus adapté que le modèle NMO, car plus à même de prendre en compte les non-linéarités de l'inversion à grands déports. En effet, seul le modèle NMO présente une imprécision cinématique supérieure à une demi-période d'un signal à 3.5 Hz (qui est la première fréquence inversée), pouvant conduire à des problèmes de sauts de phase lors de l'inversion. La comparaison entre une inversion des formes d'ondes anisotrope, où seule la vitesse verticale est mise à jour, et l'inversion isotrope, où je teste différents modèles initiaux et préconditionnement des données, m'a permis de définir la nature des vitesses isotropes reconstruites en fonction de l'éclairage fourni par le dispositif d'acquisition de surface. J'ai pu ainsi mettre en évidence la reconstruction des vitesses horizontales dans la partie supérieure du modèle où les arrivées grands angles ont un poids dominant dans la reconstruction, avec une évolution vers la reconstruction des vitesses verticales et NMO en profondeur au fur et à mesure que l'éclairage angulaire devient plus étroit. La reconstruction des vitesses horizontales rapides dans la partie supérieure du milieu peut être accommodée artificiellement en profondeur par des vitesses sous-estimées (inférieures à la vitesse verticale) ou par un approfondissement des réflecteurs, nécessaires à l'ajustement des temps d'arrivées des réflexions profondes. À la fin de l'article, je propose d'argumenter les différentes interprétations à partir d'un protocole de contrôle qualité des modèles, fondé sur la migration par renversement temporel, la modélisation sismique et l'estimation des ondelettes sources. Une annexe de l'article illustre l'importance du préconditionnement de l'inversion fourni par une approximation diagonale du Hessien pour distribuer de manière équilibrée les perturbations de vitesse dans le modèle.

**On the footprint of anisotropy on isotropic full waveform inversion :
the Valhall case study**

V. Prioux, R. Brossier, Y. Gholami, S. Operto, J. Virieux, O. I. Barkved and
J. H. Kommedal

Geophysical Journal International, 2011, **187**, pages 1495-1515

DOI : 10.1111/j.1365-246X.2011.05209.x, 2011.

3.2.2 Summary

The validity of isotropic approximation to perform acoustic full waveform inversion (FWI) of real wide-aperture anisotropic data can be questioned due to the intrinsic kinematic inconsistency between short-aperture and large-aperture components of the data. This inconsistency is mainly related to the differences between the vertical and horizontal velocities in

1. pour 'First Arrival Travel Time'

vertical-transverse isotropic (VTI) media. The footprint of VTI anisotropy on two-dimensional acoustic isotropic FWI is illustrated on a hydrophone data set of an ocean-bottom cable that was collected over the Valhall field in the North Sea. Multiscale FWI is implemented in the frequency domain by hierarchical inversions of increasing frequencies and decreasing aperture angles. The FWI models are appraised by local comparison with well information, seismic modeling, reverse-time migration (RTM) and source-wavelet estimation. A smooth initial VTI model parametrized by the vertical velocity V_{P0} and the Thomsen parameters δ and ϵ was previously developed by anisotropic reflection traveltime tomography. The normal moveout ($V_{NMO} = V_{P0}\sqrt{1+2\delta}$) and horizontal ($V_h = V_{P0}\sqrt{1+2\epsilon}$) velocity models were inferred from the anisotropic models to perform isotropic FWI. The V_{NMO} models allows for an accurate match of short-spread reflection traveltimes, while the V_h model, after updating by first-arrival traveltime tomography (FATT), allows for an accurate match of first-arrival traveltimes. Ray tracing in the velocity models shows that the first 1.5 km of the medium are sampled by both diving waves and reflections, while the deeper structure at the reservoir level is mainly controlled by short-spread reflections. Starting from the initial anisotropic model and keeping fixed δ and ϵ models, anisotropic FWI allows us to build a vertical velocity model that matches reasonably well the well-log velocities. Isotropic FWI is performed using either the NMO model or the FATT models as initial model. In both cases, horizontal velocities are mainly reconstructed in the first 1.5 km of the medium. This suggests that the wide-aperture components of the data have a dominant control on the velocity estimation at these depths. These high velocities in the upper structure lead to low values of velocity in the underlying gas layers (either equal or lower than vertical velocities of the well log), and/or a vertical stretching of the structure at the reservoir level below the gas. This bias in the gas velocities and the mispositioning in depth of the deep reflectors, also shown in the RTM images, are required to match the deep reflections in the isotropic approximation and highlight the footprint of anisotropy in the isotropic FWI of long-offset data. Despite the significant differences between the anisotropic and isotropic FWI models, each of these models produce a nearly-equivalent match of the data, that highlights the ill-posedness of acoustic anisotropic FWI. Hence, we conclude with the importance of considering anisotropy in FWI of wide-aperture data to avoid bias in the velocity reconstructions and mispositioning in depth of reflectors. Designing a suitable parametrization of the VTI acoustic FWI is a central issue to manage the ill-posedness of the FWI.

Keywords

Inverse theory, controlled source seismology, computational seismology, seismic anisotropy, wave propagation, seismic tomography

3.2.3 Introduction

Full waveform inversion (FWI) is a multiscale data-fitting approach for velocity-model building from wide-aperture/wide-azimuth acquisition geometries. Since the pioneering works on FWI in the eighties (Gauthier *et al.*, 1986; Mora, 1987, 1988; Neves & Singh, 1996), the benefits of wide apertures or long offsets to reconstruct long and intermediate wavelengths of a medium, and hence, to improve FWI resolution, have been recognized. In the framework of frequency-domain FWI, Pratt & Worthington (1990); Pratt *et al.* (1996); Pratt (1999); Sirgue & Pratt (2004) also show how the redundant wavenumber coverage provided by wide-aperture surveys

can be taken advantage of in the design of efficient FWI algorithms, as applied to decimated data-sets that correspond to a few discrete frequencies. On the other hand, wide apertures and long offsets increase the non-linearity of the inversion, because the wavefronts integrate the medium complexity over many propagated wavelengths, and then make the local optimization subject to cycle-skipping artifacts (Sirgue, 2006; Pratt, 2008). Cycle skipping artifacts will arise when the relative traveltime error, namely, the ratio between the traveltime error and the duration of the simulation, exceeds half of the inverse of the number of propagated wavelengths (Pratt, 2008; Virieux & Operto, 2009). The most efficient approach to mitigate these non-linearities consists of the recording and inverting of sufficiently-low frequencies. For typical hydrocarbon exploration surveys with a few kilometers penetration depth, maximum recording distances 10 - 20 km and maximum recording times 10-20 s, these frequencies can be as low as 1.5-2 Hz, and correspond to few propagated wavelengths (Plessix, 2009a; Plessix *et al.*, 2010a; Plessix & Perkins, 2010). These quite low frequencies can allow the FWI to be started from a crude laterally-homogeneous initial model (Plessix *et al.*, 2010a). Alternatively, several multiscale strategies have been proposed to mitigate the non-linearity of the FWI. The most usual one consists of hierarchical inversions of sub-datasets of increasing high-frequency content (Bunks *et al.*, 1995). In the frequency-domain, these sub-datasets generally correspond to a few discrete frequencies, which are inverted sequentially from the lower frequencies to the higher frequencies (e.g., Ravaut *et al.*, 2004; Sirgue & Pratt, 2004; Operto *et al.*, 2006; Jaiswal *et al.*, 2009). A second level of multiscaling can be implemented by hierarchical inversions of decreasing aperture angles through time damping (Brenders & Pratt, 2007b; Brossier *et al.*, 2009b; Shin & Cha, 2009) or double beam-forming (Brossier & Roux, 2011), or by using offset windows in the framework of layer-stripping approaches (Shipp & Singh, 2002; Wang & Rao, 2009).

Most of the recent applications of FWI to real wide-aperture data have been performed in isotropic acoustic approximations, where only the reconstruction of the P-wave velocity is sought (e.g., Ravaut *et al.*, 2004; Operto *et al.*, 2006; Bleibinhaus *et al.*, 2007; Jaiswal *et al.*, 2009). In this framework, it is possible to question the real meaning of the reconstructed velocity, and therefore, of the validity of the isotropic approximation for the inversion of wide-aperture data, which potentially contain a significant footprint of anisotropy. An analysis of this footprint is presented by Pratt & Sams (1996), who apply isotropic and anisotropic first-arrival traveltime tomography (FATT) to cross-hole data recorded in a fractured, highly-layered medium. They show the need to incorporate anisotropic effects into the tomography, to reconcile cross-hole seismic velocities with well information. A FWI case study is also presented by Pratt *et al.* (2001), who show that isotropic and anisotropic FWI of cross-hole data allow them to match the data equally well. However, the anisotropic velocity model is significantly smoother than its isotropic counterpart, which suggests some layer-induced extrinsic anisotropy in the isotropic reconstruction. More recent case studies of mono-parameter anisotropic FWI are briefly presented by Plessix & Perkins (2010) and Vigh *et al.* (2010).

In the present study, we have addressed the validity of the isotropic approximation in the framework of FWI of surface wide-aperture data through a case study of real data from the Valhall field in the North Sea. This case study clearly highlights the footprint of vertical-transverse isotropic (VTI) anisotropy on the velocity reconstruction performed by isotropic FWI of the wide-aperture data.

The Valhall oilfield in the North Sea is characterized by the presence of gas, which hampers the imaging of the reflectors at the oil-reservoir level, and by a significant VTI anisotropy

(Kommedal *et al.*, 2004). Three-dimensional (3D) acoustic isotropic FWI has been applied to ocean-bottom cable (OBC) data by Sirgue *et al.* (2009, 2010). The resulting velocity model shows a series of complex channels at a depth of 150 *m*, and it reaches a resolution that allows it to distinguish details like fractures filled with gas at a depth of 1000 *m* (Fig. 3.1b). Using the FWI velocity model as the background model improves the migrated images of the reservoir at around a depth of 2500 *m*, and the overburden. Although the 3D FWI shows impressive results, Sirgue *et al.* (2010) question the meaning of the isotropic velocities, as anisotropy is well acknowledged in the Valhall zone. The isotropic approximation should be acceptable if the medium is quasi-elliptic (low values of the η parameter (Thomsen, 1986) and if the lateral velocity contrasts are smooth enough. If these conditions are satisfied, the main effects of the anisotropy in the isotropic FWI models should be a vertical stretching of the velocity structure. The meaning of the isotropic velocity reconstructed from anisotropic wide-aperture data needs, however, to be clarified. These isotropic velocities can be steered towards horizontal, normal-moveout or vertical velocities depending on the local angular coverage provided by the acquisition geometry. This is the question we want to address in this present study by comparing the FWI velocity models inferred by 2D anisotropic and isotropic FWI of wide-aperture data. For this, we design a complete FWI workflow with *a-posteriori* model appraisal based upon reverse time migration (RTM), seismic modeling and source-wavelet estimation.

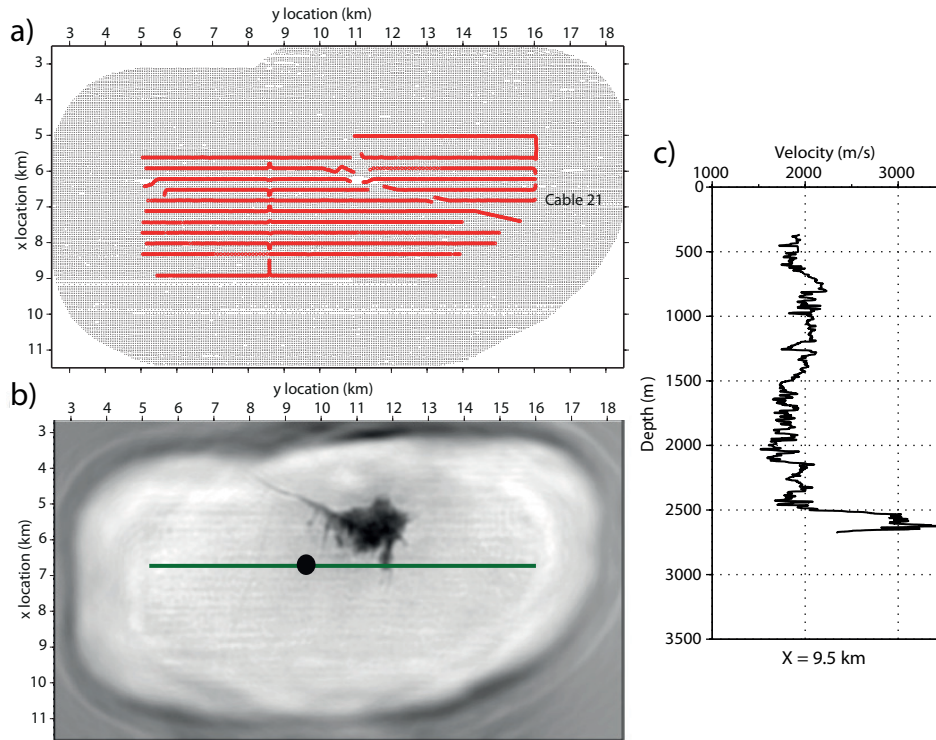


FIGURE 3.1 – The Valhall experiment - (a) Layout of the Valhall survey. Points and lines denote the positions of shots and the 4C-OBC, respectively. Cable 21 is the 2D line considered in the present study. (b) Horizontal slice at a depth of 1000 *m* across the gas cloud extracted from the 3D-FWI model of Sirgue *et al.* (2010) (from Sirgue *et al.* (2009)). The black line matches the position of cable 21. The black circle gives the position of the well log. (c) Vertical-velocity well log extracted at position $x = 9500$ *m* (courtesy of BP).

In the next section, we briefly review the main features of the FWI that we use. Then, we present the application of this imaging technique to the Valhall data set. Here, we first provide an interpretation of the main arrivals recorded in the OBC receiver gathers. Second, we discuss the accuracy of the possible initial velocity models for isotropic FWI of wide-aperture data; namely, the normal moveout (NMO) and horizontal velocity models. Then, we present the results of the anisotropic and isotropic FWI, where in the latter case, different initial models and data preconditioning are tested. Some bias in the isotropic FWI models is discussed based on the comparison of the FWI results with a vertical-seismic-profiling (VSP) log that is available for the profile, and the results of anisotropic and isotropic RTM. Before concluding, we discuss the ill-posedness of FWI, in terms of the non-unicity of the solution by comparing the data fit and the source wavelets obtained by anisotropic and isotropic FWI.

3.2.4 Methods

In the present study, isotropic and anisotropic acoustic FWI is performed in the frequency domain using the elastic FWI method described by Brossier (2011b). The modeling engine was extended to VTI media by Brossier *et al.* (2010b). The VTI acoustic approximation is considered by setting the shear-wave velocity on the symmetry axis to zero and the pressure wavefield is approximated by the average of the normal stresses (Brossier *et al.*, 2010b). The seismic modeling is performed in the frequency domain with a velocity-stress discontinuous Galerkin (DG) method on unstructured triangular mesh, which allows for accurate positioning of the sources and receivers, and accurate parametrization of the bathymetry in the framework of the shallow-water environment of Valhall (Brossier *et al.*, 2008b; Brossier, 2011b). In the frequency domain, seismic modeling can be recast in matrix form as

$$\mathbf{A}\mathbf{u} = \mathbf{s}, \quad (3.1)$$

where \mathbf{A} is the sparse impedance matrix, which is also known as the forward problem operator. This depends on the frequency, the mesh geometry, the DG interpolation order of each cell, and the physical properties. The monochromatic wavefield vector is denoted by \mathbf{u} and contains the pressure and particle velocity components at each degree of freedom of the mesh. The source vector is denoted by \mathbf{s} . We solve Eq. 3.1 with the Multifrontal Massively Parallel Sparse (MUMPS) direct solver (MUMPS-team, 2009).

The inverse problem is recast as a local optimization where a norm of the data residual vector $\Delta\mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})$ in the vicinity of an initial model should be minimized iteratively. The vectors \mathbf{d}_{obs} and $\mathbf{d}_{cal}(\mathbf{m})$ denote the observed and modeled data, respectively, where $\mathbf{d}_{cal}(\mathbf{m}) = \mathcal{S} \mathbf{u}(\mathbf{m})$, and the restriction operator \mathcal{S} extracts the values of the modeled wavefield \mathbf{u} at the receiver positions.

In the present study, the misfit function is defined by the weighted least-absolute-value (L_1) norm of the data residual vector. We choose the L_1 norm in the data space because it has been shown to be less sensitive to noise in the framework of efficient frequency-domain FWI (Brossier *et al.*, 2010c), leading us to the following definition of the misfit function,

$$\mathcal{C} = \sum_{i=1,N} |s_{d_i} \Delta d_i|, \quad (3.2)$$

where $|x| = (xx^*)^{1/2}$ and N is the dimension of the data residual vector. In Eq. 3.2, the coefficients s_{d_i} of a diagonal weighting operator \mathbf{W}_d controls the relative weight of each element

of the data residual vector. The updated model at iteration $(n+1)$ is related to the initial model (i.e., the final model of iteration (n)) and to the perturbation model $\delta\mathbf{m}^{(n)}$ by

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \alpha^{(n)}\delta\mathbf{m}^{(n)}, \quad (3.3)$$

where $\alpha^{(n)}$ denotes the step length estimated by line search. Minimization of the misfit function, Eq. 3.2, leads to the following expression of the perturbation model $\delta\mathbf{m}$

$$\delta\mathbf{m}^{(n)} = -\mathbf{B}^{(n)^{-1}}\nabla C^{(n)}, \quad (3.4)$$

where the operator $\mathbf{B}^{(n)}$ denotes the Hessian matrix (e.g., Tarantola, 2005). In the present study, we shall use only the diagonal terms of the so-called approximate Hessian matrix (i.e., the linear part of the full Hessian) damped by a prewhitening factor (Ravaut *et al.*, 2004, their equation 15), as a preconditioner of the Polak & Ribière (1969) preconditioned conjugate-gradient method, where the diagonal terms of the approximate Hessian are aimed at correcting for geometrical spreading of the data residuals and the partial derivative wavefields. Moreover, the descent direction is steered towards smooth models by filtering out the high-wavenumber components of the gradient by 2D Gaussian smoothing (e.g., Sirgue & Pratt, 2004; Ravaut *et al.*, 2004; Guitton *et al.*, 2010). The gradient ∇C of the misfit function C is computed using the adjoint-state method (Plessix, 2006), which gives the following expression for the gradient,

$$\nabla C_{m_i} = -\Re\left\{\mathbf{u}^t \frac{\partial \mathbf{A}^T}{\partial m_i} \bar{\lambda}\right\}, \quad (3.5)$$

where the real part of a complex number is denoted by \Re , the conjugate of a complex number by the sign $-$, and the so-called adjoint wavefield by λ . In Eq. 3.5, the gradient is given for one frequency and one source. The gradient that corresponds to multiple sources and frequencies is computed as the sum of the elementary gradients associated with each source-frequency couple. For the L_1 norm, the adjoint wavefield is computed by back-propagating the weighted data residuals that are normalized by their modulus (Brossier *et al.*, 2010c),

$$\mathbf{A}\bar{\lambda} = \mathcal{S}^t \bar{\mathbf{r}}, \quad (3.6)$$

where $\bar{\mathbf{r}}_i = s_{d_i} \bar{\Delta d_i} / |\Delta d_i|$. The operator $\partial \mathbf{A} / \partial m_i$, Eq. 3.5, describes the radiation pattern of the virtual secondary source of the partial derivative wavefield with respect to the model parameter m_i (Pratt *et al.*, 1998).

To increase the quadratic-well-posedness of the inverse problem (Chavent, 2009, p. 162), the FWI algorithm is designed into a multiscale reconstruction of the targeted medium (Brossier *et al.*, 2009b; Brossier, 2011b). The first level of multiscaling is controlled by the outer loop over the frequency groups, where a frequency group defines a subset of simultaneously-inverted frequencies. The multiscale algorithm proceeds over frequency groups of higher-frequency content, with possible overlap between frequency groups. A second level of multiscaling is implemented within a second loop over exponential time-damping applied from the first-arrival times t_0 . The time damping is implemented in the frequency domain by means of complex-valued frequencies, where the imaginary part of the frequency controls the amount of damping (Brenders & Pratt, 2007b; Brossier *et al.*, 2009b; Shin & Cha, 2009). A damped wavefield \mathbf{u} can be written in the frequency domain as

$$\mathbf{u}\left(\omega + \frac{i}{\tau}\right)e^{\frac{t_0}{\tau}} = \int_{-\infty}^{+\infty} \mathbf{u}(t)e^{-\frac{(t-t_0)}{\tau}} e^{i\omega t} dt, \quad (3.7)$$

where τ is the time-damping factor (s). The time-damping preconditioning injects in progressively more data during one frequency-group inversion : specifically shorter-aperture seismic arrivals are progressively involved as the time-damping factor τ increases. During the early stages of the frequency-group inversion, the early-arriving phases are mainly used to favor the long-wavelength reconstructions in the framework of the multiscale imaging (Sheng *et al.*, 2006). Frequency-domain FWI algorithms based upon the two loops over the real part and the imaginary parts of the frequency domain were also referred to as Fourier-Laplace inversion by Shin & Cha (2009).

In real-data application, the source-wavelet signature $s(\omega)$ is generally unknown, and so it must be estimated for each frequency. As the source is linearly related to the wavefield (see Eq. 3.1), the source-wavelet signature can be estimated by solving a least-squares linear inverse problem, assuming that the medium is known. Following Pratt (1999), we reconstructed the source function s in the frequency domain through the expression

$$s(\omega) = \frac{\mathbf{g}_{cal}(\omega)^T \mathbf{d}_{obs}^*(\omega)}{\mathbf{g}_{cal}(\omega)^T \mathbf{g}_{cal}^*(\omega)}, \quad (3.8)$$

where \mathbf{g}_{cal} denotes the Green functions at the receiver positions through the relationship $\mathbf{d}_{cal}(\omega) = s(\omega)\mathbf{g}_{cal}(\omega)$. In the framework of the adjoint-state method, for consistency with the model update performed with an L_1 norm minimization, the source signature can also be estimated alternatively with an L_1 norm minimization (personal communication, R. -E. Plessix, 2010). Such optimization has been implemented with a non-linear optimization scheme based on the very fast simulated annealing (VFSA). Our experience with source-wavelet estimation shows that for both synthetic and real data sets, non-linear L_1 and linear L_2 optimizations give similar results. The source signatures are updated for each source gather at each iteration once the incident Green functions \mathbf{g}_{cal} are computed, and they are subsequently used for the gradient computation and model update.

3.2.5 Application to Valhall

3.2.5.1 Geological context and acquisition geometry

Geological context

The Valhall oilfield in the North Sea has been producing oil since 1982. This is a shallow-water environment (water depth 70 m) that is located in the central zone of an old Triassic graben, that entered into compression during the late Cretaceous (Munns, 1985). The subsequent inversion of stress orientations led to the formation of an anticlinal that now lies at a depth of 2.5 km, creating a high-velocity contrast respect to overlying layers. An extension regime occurred in the tertiary age, allowing for a thick deposit of sediments with gas trapped in some layers. In rising from the underlying Jurassic layers, oil was trapped underneath the cap rock of the anticlinal. The oil migration reaches a peak nowadays, by means of numerous fractures that were induced by the different tectonic phases. Of note, these fluid are the cause of the high porosity preservation of the Valhall reservoir : a distinctive feature even though this field is affected by subsidence, which is likely to be due to production.

Ocean-bottom-cable (OBC) acquisition geometry and initial models

The layout of the 3D wide-aperture/azimuth acquisition designed by the company BP is shown in Fig. 3.1a, where the black points and the lines represent the locations on the sea floor of the shots at 5 *m* depth and of the permanent OBC-four-component arrays at around 70 *m* depth, respectively. One cable contains 220 4-C receivers. In the present study, 2-D acoustic FWI is applied to the OBC line indicated as cable 21 in Fig. 3.1a. This line corresponds to 320 shots recorded by 220 4-C receivers for a maximum offset of 13 km. This cable is located outside the gas cloud, as shown on the horizontal cross-section that was extracted at a depth of 1000 *m* from the 3D FWI model of Sirgue *et al.* (2010) (Fig. 3.1b). A VSP log for vertical velocity is available on line 21 and it will be used to locally assess the FWI results (Fig. 3.1c). A low-velocity zone that results from the presence of gas layers is clearly seen on the VSP log, between 1.5 *km* and 2.5 *km* in depth.

A 3D model for the vertical velocity V_{P0} and the Thomsen parameters δ and ϵ (Thomsen, 1986) has been developed by anisotropic reflection traveltime tomography in VTI media, and is provided by BP (Fig. 3.2(a-c)). The vertical velocity model shows the low-velocity zone associated with the gas layers between 1.5 km and 2.5 km in depth, above the reservoir level (Fig. 3.2a). The corresponding normal moveout (NMO) and horizontal velocity models are shown in Fig. 3.2(f-g). In this study, by NMO velocity is meant the wave speed given by $V_{NMO} = V_{P0}\sqrt{1+2\delta}$ (Tsvankin, 1995), while the horizontal velocity is given by $V_h = V_{P0}\sqrt{1+2\epsilon} = V_{NMO}\sqrt{1+2\eta}$, respectively, where the anellipticity coefficient η is given by $\eta = (\epsilon - \delta)/(1 + 2\delta)$ (Alkhalifah & Tsvankin, 1995). The NMO velocities should allow the short-spread reflection traveltimes in VTI media to be matched (Tsvankin, 2001), while the horizontal velocities should allow the refraction and long-spread reflection traveltimes to be matched. Both velocity models can be viewed as initial models of isotropic FWI of wide-aperture seismic data, as both short-aperture reflections and diving waves are recorded by long-offset acquisition and are involved in this FWI processing. The 3D FWI model developed by Sirgue *et al.* (2010) is obtained using the NMO velocity model as the initial model (L. Sirgue, personal communication, 2010). The percentage of the anisotropy in Valhall is shown by $\eta \approx (V_h - V_{NMO})/(V_{NMO})$, and it reaches a maximum value of 16 % (Fig. 3.2d).

3.2.5.2 Anatomy of the data and starting-model appraisal

A receiver gather for line 21 is shown in Fig. 3.3a. The main phases that can be interpreted are : (1) The first arrivals (Fig. 3.3a, D1, D2). The traveltime curve in the offset-time domain shows two distinct slopes with a crossover distance of around 4000 *m*, which suggests the presence of an interface in the upper structure, with the reflection from this interface shown in Fig. 3.3a (Rs). (2) The reflection from the top of the gas layers (Fig. 3.3a, Rg), which can be followed at long offsets where the traveltime curve becomes tangential to the D2 phase. (3) The reflection from the top of the reservoir at the base of the gas layers (Fig. 3.3a, Rr). The reflections from the top of the reservoir are disrupted at critical and super-critical distances by shingling dispersive guided waves propagated in the near surface (Fig. 3.3a, SW) (Robertson *et al.*, 1996). These high-amplitude waves can have a harmful impact on the acoustic inversion procedure, in particular, if the sea bottom is not accurately modeled, and because they contained PS converted waves which are not accounted for by the acoustic modeling.

To validate our former interpretation and assess the kinematic accuracy of the NMO and horizontal velocity models, we compute the first-arrival traveltimes and reflection traveltimes

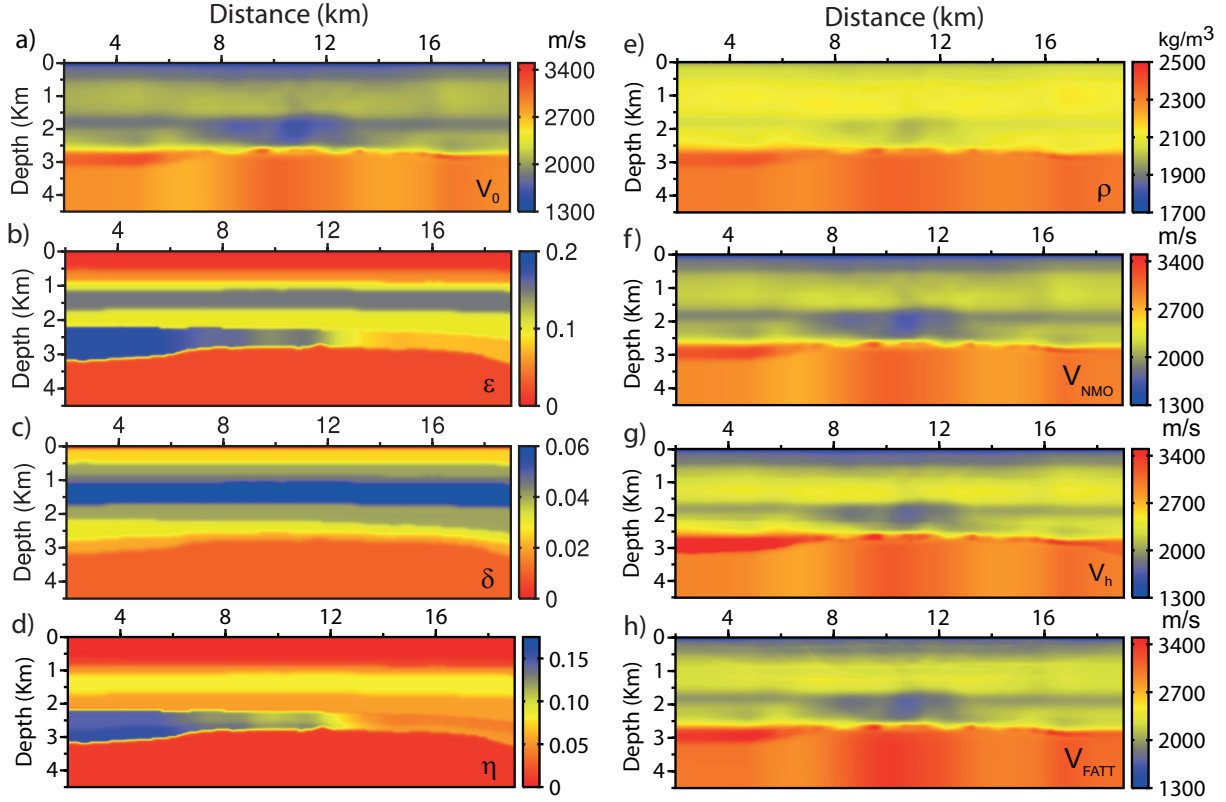


FIGURE 3.2 – Two-dimensional sections along position of cable 21 through anisotropic 3D models of the Valhall field. (a) Vertical velocity (V_{P0}). (b) Thomsen parameter ϵ . (c) Thomsen parameter δ . (d) Anellipticity parameter η . (e) Density ρ . (f) NMO velocity (V_{NMO}). (g) Horizontal velocity (V_h). (h) Horizontal velocity updated by the first-arrival traveltimes tomography (FATT model). The V_{P0} , δ , and ϵ models were built by reflection tomography (courtesy of BP). The density model was inferred from the V_{NMO} model using the Gardner law.

from the top and the bottom of the gas layers using the isotropic eikonal solver of Podvin & Lecomte (1991) (Fig. 3.3(b-c)). We roughly manually pick the top and the bottom of the gas layers to compute the reflection traveltimes. For a maximum offset of 11 km, the rays associated with the first-arrival traveltimes turn at a maximum depth of 1.5 km, and, therefore, they do not sample the structure at the reservoir level below the gas layers. This implies that the first 1.5-km of the structure are constrained by both diving waves and reflected waves, while the deeper structure is mostly constrained by short-spread reflected waves. Superimposition of the computed traveltimes curves on the receiver gather shows that the NMO velocities do not allow the traveltimes at long offsets of diving waves and the long-spread reflection Rg to be matched (Fig. 3.3b). The mismatch between the observed and computed first-arrival traveltimes reaches around 0.3 s at 11 km of offset. Cycle skipping artifacts will occur when this traveltimes error exceeds half the period of the signal, *i.e.* for a frequency as low as roughly 1.7 Hz. In the following, we use an initial frequency of 3.5 Hz for inversion, which allows for a maximum traveltimes error of 0.14 s, which is reached for an offset of the order of 6 km. We conclude that FWI might be affected by cycle skipping artifacts that result from the inversion of the diving waves and long-spread reflection recorded at offsets greater than 6 km, when the NMO velocity

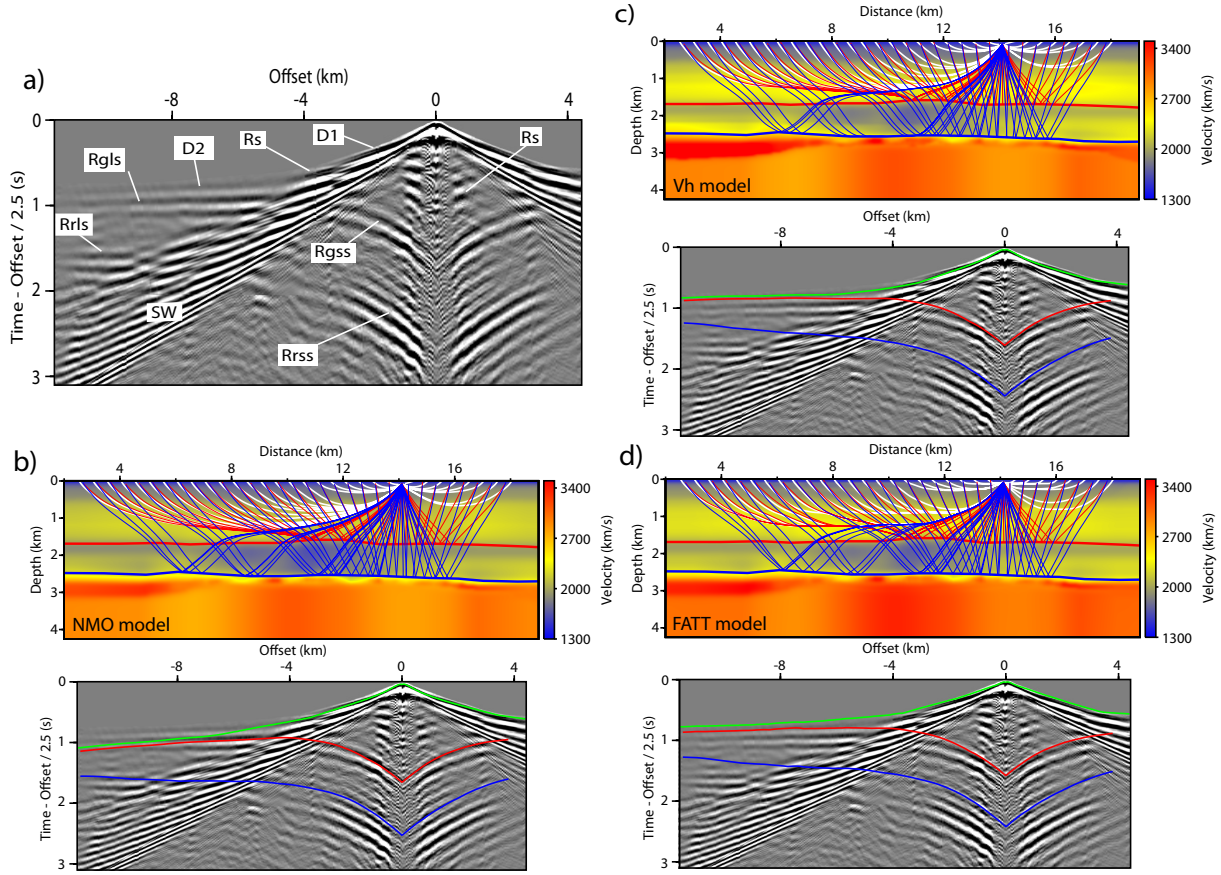


FIGURE 3.3 – OBC data set. (a) Example of preprocessed recorded receiver gather, at position $x = 14100$ m. The vertical axis is plotted with a reduction velocity of 2.5 km/s. Phase nomenclature : D1, D2 : diving waves. Rs : shallow reflection. Rgss / Rgls : short-spread and long-spread reflections from the top of the gas. Rrrs / Rrrs : short-spread and long-spread reflections from the top of the reservoir. SW : shingling waves. (b) Top : ray tracing in the NMO velocity model for the first arrival (white rays), and the reflections from the top of the gas (red) and the reservoir (blue). Top of the gas and the reservoir are delineated by red and blue solid lines, respectively. Bottom : receiver gather shown in (a) with superimposed traveltime curves computed in the NMO model for these three phases. (c) As (b) for the V_h model. (d) As (c) for the FATT model. See text for details.

model is used as an initial model. In contrast, the NMO velocity model is expected to make the short-spread reflection traveltimes of the phases Rg and Rr to be matched, that is supported by Fig. 3.3b. Unlike the Rg phase, the NMO velocity model reasonably predicts the slope of the long-spread reflection traveltimes of the Rr phase. Indeed, the reflection-angle illumination of the reflectors decreases with depth, which should make the reflection traveltime curve associated with the top of the reservoir less sensitive to anisotropy (i.e., the difference between vertical and horizontal velocities) within the recorded offset range. The horizontal velocity model allows for a much better agreement of the first-arrival traveltimes (Fig. 3.3c). The reflection traveltimes of the phase Rr computed in this model are lower than the traveltimes computed in the NMO model. The mismatch between the NMO and horizontal velocity reflection traveltimes at zero

offset is of the order of 0.075 s and 0.1 s for the Rg and Rr phases, respectively. Assuming that the NMO traveltimes accurately predict the observed reflection traveltimes, it is worth mentioning that these traveltime mismatches remain below the cycle-skipping limit of 0.14 s because short-offset reflection data involve fewer propagated wavelengths than long-spread reflections and diving waves. We might conclude from this analysis that the horizontal velocity model should provide a more suitable initial model than the NMO velocity model for isotropic FWI because the traveltime errors remains always below the cycle-skipping limit whatever the offsets. However, the NMO velocity model is expected to provide the most accurate match of the short-aperture reflection traveltimes.

The horizontal-velocity model does not accurately match the first-arrival traveltimes at intermediate offsets (with a maximum error of the order of 0.1 s at 5.5-km offset). This highlights that seismic reflection data are not suitable for accurate reconstruction of horizontal velocities. This prompted us to update the horizontal-velocity model by FATT to improve the match of the first-arrival traveltimes before FWI (Figs 3.2h and 3.3d). This updated velocity model will be referred to as the FATT model in what follows.

In continuing this study, we use the NMO model and the FATT models as initial models for isotropic FWI, and we compare the isotropic FWI models with the results of anisotropic FWI for vertical velocity.

3.2.5.3 FWI preprocessing and experimental setup

FWI preprocessing

Among the available 4-C receiver data, only the hydrophone component is considered as we are dealing with acoustic FWI. Acoustic inversion was applied to the hydrophone component of the fully elastic data computed in the synthetic elastic Valhall model (Brossier *et al.*, 2009c) : a successful image of the V_P structure has been obtained because converted P-SV waves have a minor footprint on the hydrophone component. Therefore, Valhall should provide a suitable framework for the successful application of acoustic FWI to elastic data (see Barnes & Charara (2009) for a more general discussion on the validity of the acoustic approximation in the marine environment). As the receivers are around 2/3-fold less numerous than the shots, the data are sorted in receiver gathers by virtue of the source-receiver reciprocity holding between an explosion source and a pressure component of the data, to reduce the computational cost.

The FWI data preprocessing first consists of minimum-phase whitening followed by Butterworth filtering of a [4–20] Hz bandwidth. The whitening is designed to preserve the geometrical spreading of the data, by normalizing the spectral amplitudes of the deconvolution operator associated with each trace according to its maximum amplitude. The bandwidth of the Butterworth filter is chosen heuristically to provide the best trade-off between the signal-to-noise ratio and the flattening of the amplitude spectrum. We then apply FK filtering to remove as much S-wave energy as possible, and spectral matrix filtering (Mari *et al.*, 1999, page 386) to enhance the lateral coherency of events (Ravaut *et al.*, 2004). We also applied a mute to remove noise before the first-arrival time, and after a time of 4 s following the first-arrival excluding late arrivals. Finally, the data are multiplied by the function \sqrt{t} to roughly transform the 3D geometrical spreading of real amplitude data into a 2D amplitude behaviour. An example of a fully preprocessed receiver gather is shown in Fig. 3.3a.

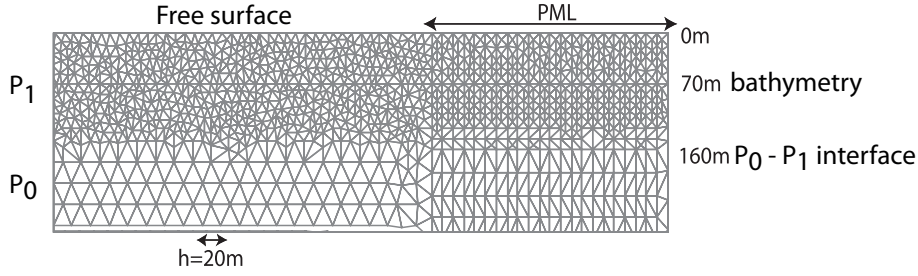


FIGURE 3.4 – Seismic modeling - Close-up of the hybrid P1-P0 triangular mesh on which seismic modeling was performed using the DG method.

Experimental set-up : seismic modeling

The $18000\text{ m} \times 5000\text{ m}$ velocity, density and attenuation models are discretised on unstructured triangular meshes for seismic modeling with the DG method, where the medium properties are piecewise constant per element (Brossier *et al.*, 2008b; Brossier, 2011b). Accurate positioning of the seismic devices is allowed by the use of a fine mesh in the first 160 m of the medium, where the linear interpolation order (P1) is used to describe the acoustic wavefield (Fig. 3.4). Below, a regular triangular mesh is used with piecewise-constant (P0) representation of the wavefield in each cell to reduce the cost of the modelling in terms of memory and computation. A discretisation rule of 10 elements per wavelength is used in the regular mesh, that leads to 20-m-long triangle edges. The hybrid P1-P0 mesh contains around 585×10^3 cells. The mesh includes 500-m-thick perfectly matched layers on the right, left and bottom sides of the model for the absorbing boundary conditions (Berenger, 1994). A free-surface boundary condition is implemented on top of the models, which implies that free-surface multiples are involved during the FWI. Although the real depth of the receivers varies between 67 m and 73 m, we choose for convenience the design of a flat bathymetry at a depth of 70 m within the mesh : all receivers are put at a depth of 71 m, just below the sea bottom. This approximation has a minor impact on the modeling accuracy given the shortest propagated wavelength of 215 m.

Experimental set-up : inversion

Only the P-wave velocity is reconstructed during the inversion procedure we perform. An attenuation model is set as homogeneous below the sea bottom to the realistic value of the attenuation factor $Q_p = 150$. This value of attenuation is chosen by trial-and-error, such that the root-mean-squares amplitudes of the early-arriving phases computed in the initial model roughly matches those of the recorded data, following the approach of Pratt (1999, his figure 6). The density model is inferred from the starting FWI velocity models using the Gardner law (Gardner *et al.*, 1974) and is kept constant over iterations of the inversion (Fig. 3.2e).

We sequentially invert five increasing frequency groups between 3.5 Hz and 6.7 Hz ([3.5, 3.78, 4], [4, 4.3, 4.76], [4.76, 5, 5.25], [5.25, 5.6, 6], [6, 6.35, 6.7] Hz). We have verified that a sufficiently high signal-to-noise ratio is inside the traces at the lowest frequency of 3.5 Hz, as already used in the 3D FWI application of Sirgue *et al.* (2010). The spectral amplitude of the 3.5-Hz frequency represents 45 % of that of the dominant 7-Hz frequency after whitening and Butterworth filtering. The maximum frequency of 6.7 Hz is similar to that used in Sirgue *et al.* (2010). We do not investigate yet whether the FWI can be pushed towards higher frequencies for this case study. We design our frequency groups with three frequencies per group, with

one-frequency overlapping between the groups.

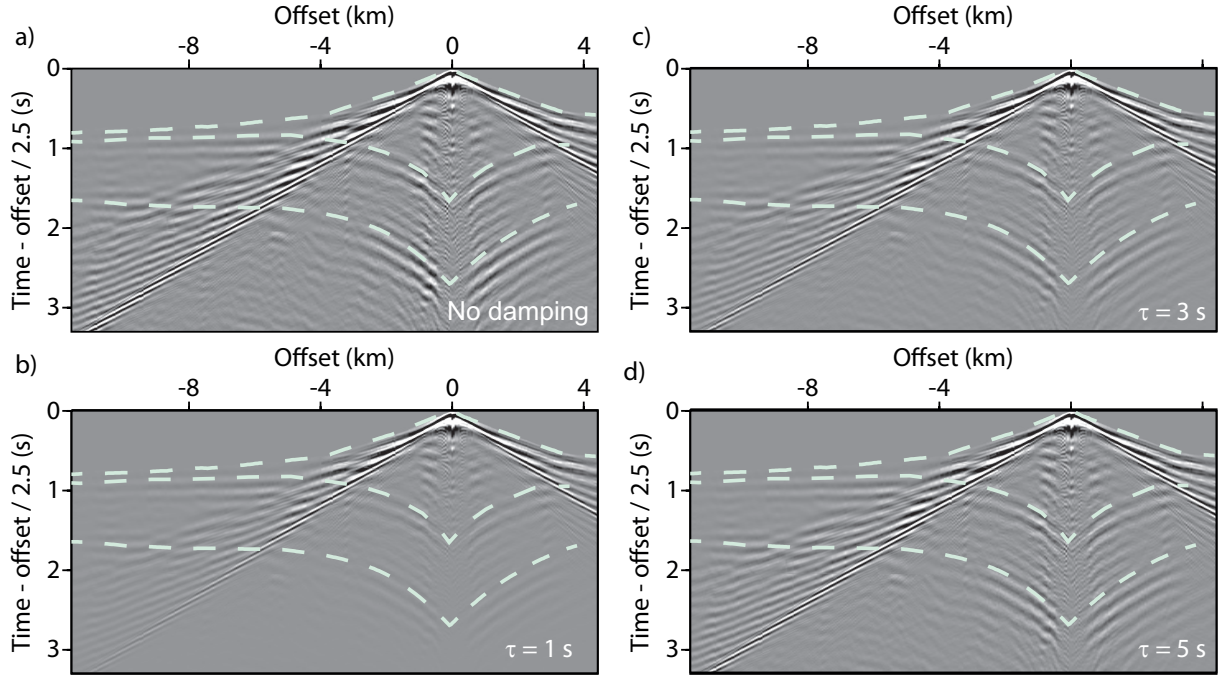


FIGURE 3.5 – Data preconditioning according to time damping - Recorded receiver gathers plotted without time damping (a), and with time damping using $\tau = 1$ s (b), 3 s (c) and 5 s (d). The dashed lines show the first-arrival picks and the reflection picks from the top and the bottom of the gas layers.

When used, the time-damping factors τ are chosen as 1 s, 3 s and 5 s, and were applied from the first-arrival travel time t_0 . Fig. 3.5 shows a receiver gather where we applied various time-damping factors. When the most aggressive damping factor ($\tau = 1$ s) is applied during the early iterations of one frequency-group inversion, the inversion favors the early-arriving phases that are associated with the wide-aperture components of the data, while the use of higher values of τ progressively introduces later-arriving phases that are associated with shorter-aperture components.

We allow a maximum of 25 iterations within the nested loops over frequency groups and time damping (Brossier *et al.*, 2009b). We note that 25 iterations are enough, because the preconditioned conjugate-gradient algorithm often stopped before reaching this step as we stop the iterations as soon as the maximum velocity perturbation is lower than 10^{-3} % of the velocity of the starting model at the position of the maximum velocity perturbation.

As the data are sorted by receiver gather for FWI, we estimate a source wavelet per receiver gather at each non-linear FWI iteration. The underlying assumption is that the shots are perfectly repetitive.

Gaussian smoothing is applied to the model velocity perturbations using horizontal and vertical correlation lengths of $0.3 \times \lambda$ and $0.1 \times \lambda$, respectively, where λ is the local wavelength. Of note, the horizontal correlation length is set three-times longer than the vertical one, as the medium is fairly tabular. The model is kept constant down to a depth of 77 m to keep the velocity in the water layer constant (above 70 m depth), and to avoid instabilities in the vicinity

of the sources and receivers located at 71-m and 5-m depth, respectively. No data weighting is applied during FWI, and, therefore, $\mathbf{W}_d = \mathbf{I}$, where \mathbf{I} is the identity matrix.

As reported by Pratt & Shipp (1999), the potential improvement in resolution provided by FWI compared to FATT, can be estimated as being of the order of $\sqrt{N_\lambda}$, where N_λ is the number of wavelengths propagating between the source and the receiver. Although we also consider starting models originally built by reflection traveltime tomography, that shows higher resolution than models built by FATT only, it might be interesting to quantify this resolution improvement. The velocities in the Valhall model range between 1500 m/s and 3500 m/s , and FWI is performed in the [3.5-7] Hz frequency band. For a maximum offset of 13 km , it follows that FWI should lead to an increase in resolution by a factor between 4 and 8 compared to the resolution of FATT.

3.2.5.4 FWI results

We perform five applications of FWI, for which we use different approximation (isotropic *versus* anisotropic), starting models $((V_{P0}, \delta, \epsilon), \text{NMO}, \text{FATT models})$, and data preconditioning (without and with time dampings). The main features of the five inversion applications are outlined in Table 3.1.

FWI test	Anisotropic / Isotropic FWI	Initial FWI model \mathbf{m}_0	τ (s)
V_{P0} +FWI	Anisotropic	$(V_{P0}, \delta, \epsilon)$	-
NMO+FWI	Isotropic	NMO	-
FATT+FWI 1	Isotropic	FATT	-
FATT+FWI 2	Isotropic	FATT	1, 3, 5 s
FATT+FWI 3	Isotropic	FATT	1 s

TABLE 3.1 – FWI setup used for the five applications in the present study. \mathbf{m}_0 : initial FWI model. $\tau(s)$: time damping. A / I : Anisotropic versus isotropic FWI. The same frequency groups are used for each of these tests (see text).

Anisotropic FWI for vertical velocity

We first perform anisotropic FWI, the results of which will be used in the following sections of this study as a reference to assess the footprint of the anisotropy on the isotropic FWI. For inversion, the VTI medium is parametrized by the vertical velocity V_{P0} and the Thomsen parameters δ and ϵ . Only the vertical velocity V_{P0} is involved in the inversion, while the δ and ϵ parameters are kept fixed during the inversion. The $(V_{P0}, \delta, \epsilon)$ FWI parametrization is justified because the vertical velocity has a dominant weight in the optimization process with respect to the Thomsen parameters. This allows for mono-parameter inversion as long as the background δ and ϵ models describe the large wavelengths of the medium accurately (Gholami *et al.*, 2011a). This latter assumption is strongly supported because the NMO and the horizontal velocity models inferred from the initial V_{P0}, δ, ϵ models (Fig. 3.2) allow the accurate matching of the short-aperture reflection and the first-arrival traveltimes, respectively (Fig. 3.3(b-c)). Another benefit of $(V_{P0}, \delta, \epsilon)$ parametrization is that the diffraction pattern of the vertical velocity is isotropic for this parametrization, and therefore, it should provide a

broadband reconstruction of the vertical velocity (Gholami *et al.*, 2011b). Note that, when the vertical velocities are updated by mono-parameter anisotropic FWI, the horizontal velocities are updated according to the coupling between the vertical and horizontal velocities through the expression : $V_h = V_{P0}\sqrt{1 + 2\epsilon}$. Other parametrization can be viewed that involve the elastic moduli (Lee *et al.*, 2010) or the NMO velocity and η (Plessix & Cao, 2011). However, a detailed discussion on the most suitable parametrization for acoustic VTI FWI is beyond the scope of this study, and this has already been discussed in Plessix & Cao (2011) and Gholami *et al.* (2011b,a).

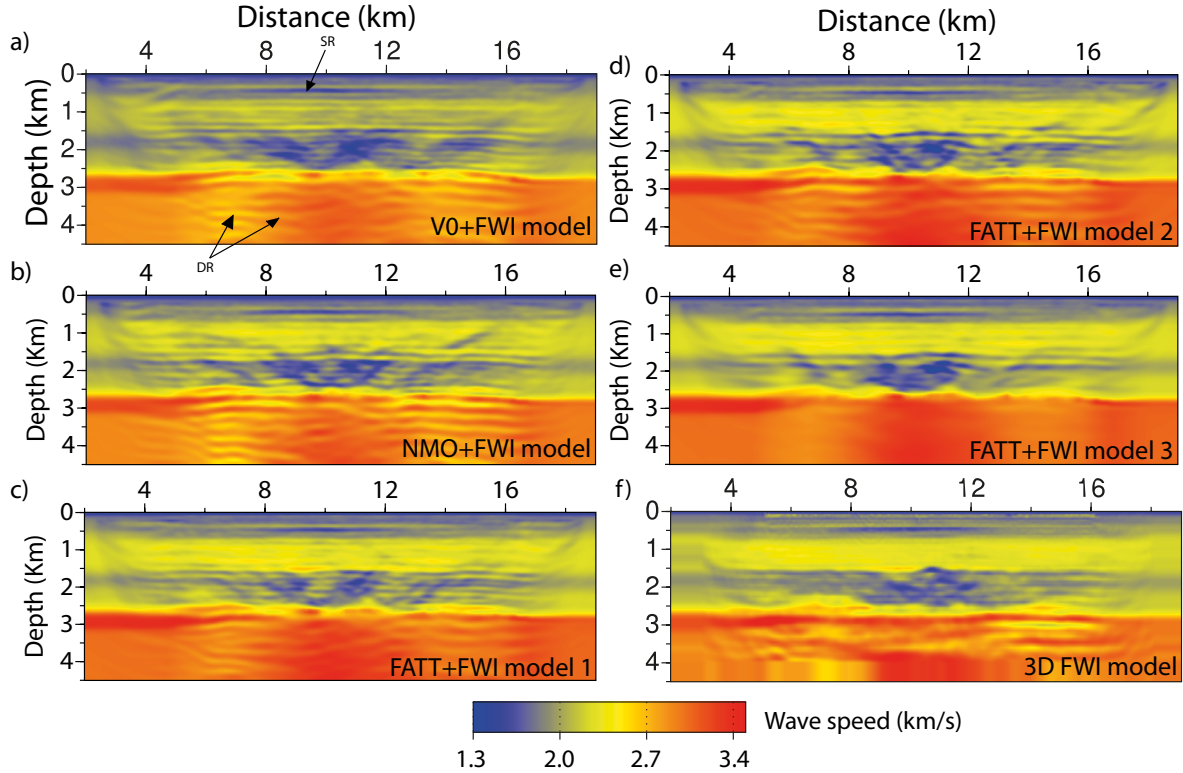


FIGURE 3.6 – FWI models. (a) V_{P0} model built by 2D anisotropic FWI. SR : shallow reflector at 0.6 km in depth described in the text. DR : deep reflector below the reservoir level described in the text. (b-e) FWI models built by 2D isotropic FWI using V_{NMO} as the initial model (b), and using V_{FATT} as the initial model without (c) and with 3 time dampings (d) and 1 time damping (e) as data preconditioning. (f) Dip section of the 3D FWI model of Sirgue *et al.* (2010) (courtesy of BP).

The initial V_{P0} model for FWI is shown in Fig. 3.2a. The background models for ϵ , δ and ρ are shown in Fig. 3.2(b,c,e). The FWI is applied consecutively to five frequency groups without considering time damping (Test $V_{P0} + FWI$ in Table 3.1). The final FWI model for V_{P0} , referred to as the $V_{P0} + FWI$ model in the following, is shown in Fig. 3.6a. A reflector at 600 m in depth is well imaged (Fig. 3.6a, SR), and might be related to the reflection phase Rs in the data (Fig. 3.3a). This reflector is consistently reconstructed for all of the tests presented below, and therefore, it should correspond to a real feature, which will not be discussed any more in the following.

The negative velocity contrasts at the top of the gas layers at 1.5 km in depth is well deli-

neated, and some deep reflectors below the reservoir level between 2.5 *km* and 4 *km* in depth can be interpreted (Fig. 3.6a, DR) (see also Sirgue *et al.*, 2010, their Figure 3b). These deep reflectors are imaged with a limited wavenumber bandwidth as suggested by their vertically-oscillating nature because of the lack of low frequencies and the lack of wide aperture coverage at these depths, which prevent the reconstruction of the small wavenumbers in the deep part of the medium. At these depths, FWI performs mostly as a least-squares migration of narrow bandwidth data. Comparisons between the log of the FWI model at 9.5 *km* in depth and the VSP log show reasonable agreement between the two velocity profiles, in particular in the upper part down to 0.8 *km* in depth (Fig. 3.7a). The log of the FWI model remains centered on the log of the initial model, which already matches the main trend of the VSP log quite well. Therefore, it is quite challenging to discriminate between the true features of the FWI log from the artificial ones. Rather than assessing the match of the $V_{P0} + FWI$ log with the VSP log in an absolute sense, we shall rather use the comparison between the $V_{P0} + FWI$ log and the VSP log as a reference to highlight the differences with the isotropic FWI results presented in the following. Of note, the VSP log is low-pass filtered in the time domain after depth-to-time conversion with a cut-off frequency of 14 *Hz*, inferred from the theoretical vertical resolution (i.e., half a wavelength) of the FWI at the 7-Hz frequency.

Isotropic FWI using the initial NMO velocity model

We use the NMO model (Fig. 3.2f) as the initial model for isotropic FWI, and we invert the full data-set without considering time damping, as for the anisotropic FWI (Test *NMO + FWI* in Table 3.1). The resulting FWI model, referred to as the *NMO + FWI* model, is shown in Fig. 3.6b. The reconstructed velocities in the upper structure are clearly higher than those of the $V_{P0} + FWI$ model, and are close to horizontal velocities between 0.6 *km* and 1.4 *km* in depth (Fig. 3.7b). The reconstruction in the gas appears more unstable than in the anisotropic case, with reconstructed velocities smaller than the vertical velocities at around 2.4 *km* in depth. The FWI successfully images deep reflectors below the reservoir level at 3 – 4 *km* in depths with higher amplitudes than for the anisotropic FWI (Compare Fig. 3.6a and 3.6b).

Isotropic FWI using the initial FATT velocity model

We consider now the same data preconditioning without time damping as for the two former tests but here we use the *FATT* model as the initial model for the isotropic FWI (Test *FATT + FWI* 1 in Table 3.1). The resulting FWI model, which is referred to as the *FATT + FWI* model 1, is shown in Fig. 3.6c. Compared to the *NMO + FWI* model, the *FATT + FWI* model 1 has slightly higher velocities between 0.6 *km* and 1.6 *km* in depth, that highlights the footprint on the initial model (Fig. 3.7c). These velocities remained centered around the horizontal velocities of the initial model at these depths. The velocities within the gas layers between 2 *km* and 2.5 *km* in depth are close to the vertical velocities along the well, and are higher than those of the *NMO + FWI* model. We note also a high-velocity perturbation at 3 *km* in depth, where the velocity reaches a maximum value of 3.3 *km/s* (Fig. 3.7b). This velocity perturbation is absent in the *NMO + FWI* model, where the maximum velocity is reached at a depth of 2.7 *km* (Fig. 3.7b). This deep velocity perturbation might indicate a vertical stretching of the deep structure to balance the high horizontal velocities reconstructed in the upper structure, and the velocities in the gas layers higher than those of the NMO model.

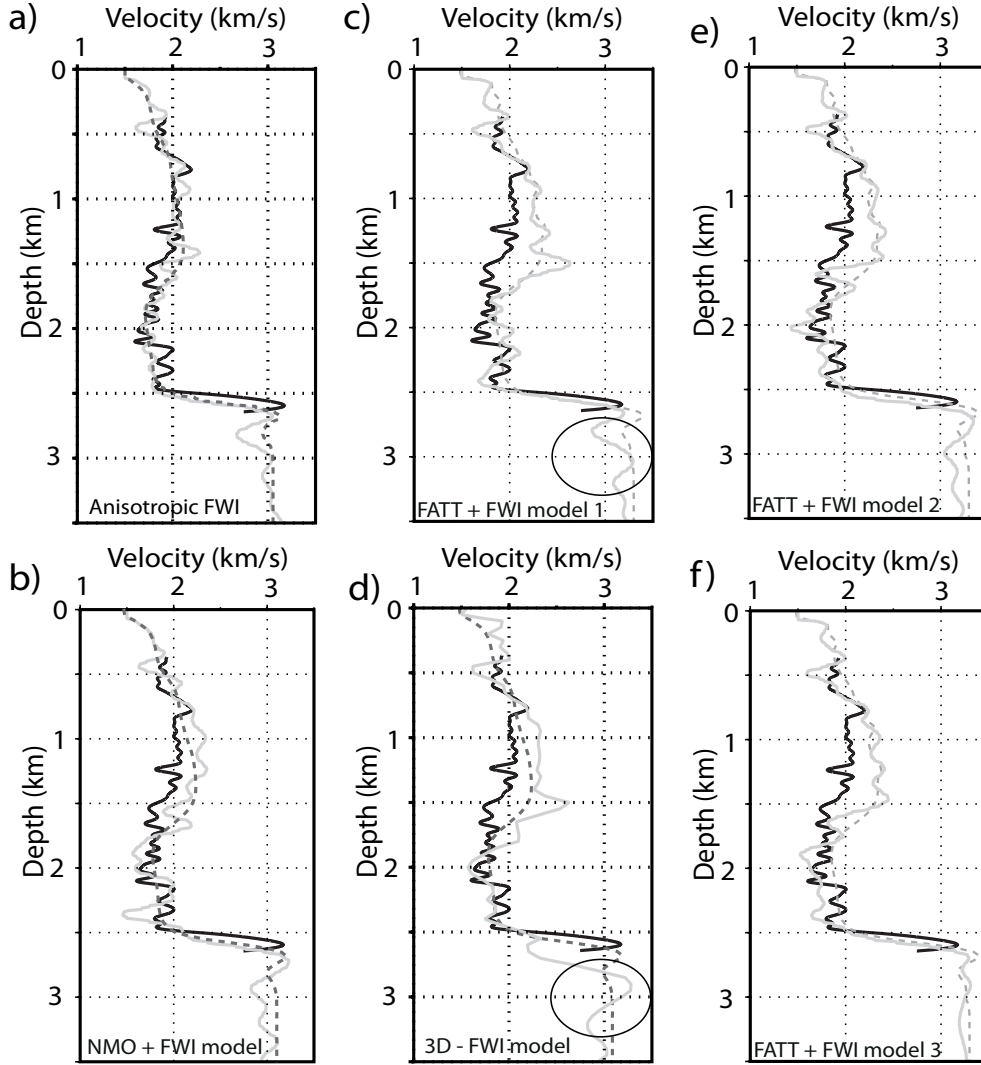


FIGURE 3.7 – Log of the initial (dash gray line) and final (solid gray line) FWI models extracted at the position of the well log. The low-pass filtered well-log vertical velocities are plotted with solid black lines. (a) Anisotropic FWI. (b,c,e,f) Isotropic FWI using V_{NMO} as the initial model (b), and using V_{FATT} as the initial model for inversion tests 3 (c), 4 (e), 5 (f) (Table 3.1). (d) The 3D FWI model of Sirgue *et al.* (2010) (courtesy of BP). In (c) and (d), the deepening of the top of the reservoir, related to the stretching in depth of the velocity model, is highlighted by the black ellipses (see text for details).

During a second test with the *FATT* model, we use three time dampings in cascade during the inversion of each frequency group ($\tau = 1, 3, 5s$) (Test *FATT + FWI* 2 in Table 3.1). Note that the wide-aperture components associated with strong time dampings are injected first during one frequency group. On one hand, this is consistent as these aperture components are those that are accurately predicted by the starting *FATT* model. On the other hand, this hierarchical strategy is consistent with the multiscale approach, where the long wavelengths constrained by the wide apertures must be first reconstructed. The final FWI model, referred to as *FATT + FWI* model 2, is shown in Fig. 3.6d. As for the *FATT + FWI* model 1, the

horizontal velocities are mainly reconstructed down to 1.5 km in depth, where the top of the gas layers is well delineated by a sharp negative velocity contrast (Fig. 3.7e). Overall, the velocities in the gas are lower than the vertical velocities along the well below 1.8 km in depth. The maximum velocity at the reservoir level is reached at 2.7 km in depth as for the NMO model.

During the third test performed with the FATT model, we consider only a time-damping factor of 1s (Test *FATT + FWI* 3 in Table 3.1). A time-damping factor of 1 s favors the aperture components of the data which are well predicted by the *FATT* model from a kinematic viewpoint, and it heavily damps the contribution of the deep short-aperture reflections in the data. The final FWI model, referred to as *FATT + FWI* model 3, is shown in Fig. 3.6e. The velocity structure of the *FATT + FWI* model 3 is similar to the one of the *FATT + FWI* model 2. However, the velocities in the gas are in overall lower and the deep part of the model is less perturbed and shows a smoother pattern due to the use of a more limited subdataset during inversion (compare Figs. 3.7e and 3.7f).

For possible identification of artifacts relating to the 3D propagation effects, we show the dip section of the 3D isotropic FWI model of Sirgue *et al.* (2010) along cable 21 (Fig. 3.6f). Interestingly, the velocities above the gas between 0.5 km and 1.5 km in depth are quite close to those of the *FATT + FWI* model 1 (compare Fig. 3.7d and 3.7c). The vertical stretching between 2.5 km and 3 km in depth of the 3D FWI model might have a similar origin than the one hypothesized for the *FATT + FWI* model 1 (compare Fig. 3.7c and 3.7d). The consistency between the dip section of the 3D FWI model of Sirgue *et al.* (2010) and the *FATT + FWI* model 1 strongly supports that 3D effects have a minor impact on the 2D FWI results.

3.2.5.5 Model appraisals

Model appraisal is a key issue in FWI as uncertainty analysis is quite challenging to perform in a Bayesian framework (Gouveia & Scales, 1998). In the present study, the FWI models are evaluated based upon four criteria : the local match with the VSP log, the flatness of the common image gathers (CIG) computed by RTM, the synthetic seismogram modeling, and the repeatability of source-wavelet estimation.

Reverse time migration and common image gathers

We compute 2D RTM and CIGs in the offset-depth domain. RTM is performed in the frequency domain using the acoustic VTI finite-difference frequency-domain modeling method of Operto *et al.* (2009) and the gradient of the FWI program of Sourbier *et al.* (2009a,b), where the data residuals are replaced by the data. Each of the common-offset migrated images were computed independently to generate CIGs before stacking. The range of offsets that is considered for migration ranges from -5 km to 5 km. For migration, we use a suitable preprocessed data set, where free-surface multiples are removed. The migrated images are displayed with an automatic gain control.

The anisotropic RTM performed in the initial $(V_{P0}, \delta, \epsilon)$ model provides a good image of the subsurface, with a good continuity of the top of the reservoir beneath the gas layers at 2.5 km in depth and of a deep reflector between 3 km and 3.5 km in depth (Fig. 3.8a, DR) (see also Fig. 3 in Sirgue *et al.* (2010)). The quality of the anisotropic migrated image is further confirmed by the overall flatness of the reflectors in the CIGs (Fig. 3.9a). The isotropic RTM

computed in the NMO model produces an acceptable image, although the image of the top of the reservoir is slightly less focused than the one obtained by anisotropic RTM (Fig. 3.8b). The deep reflector below the reservoir level is shifted downwards by around 250 *m* with respect to its position in the anisotropic image because the NMO migration velocities in isotropic RTM are on average faster than those of anisotropic RTM when short-spread data are considered. On the other hand, the reflectors are slightly smiling in the CIGs, which suggests too slow velocities at long offsets (Fig. 3.9b). The migrated image computed in the FATT model shows a severe misfocusing of the top of the reservoir with a significant deepening of the deep reflector below the reservoir level, due to the high migration velocities associated with horizontal velocities (Fig. 3.8c). These too high velocities are clearly highlighted by frowning reflectors in the CIGs (Fig. 3.9c).

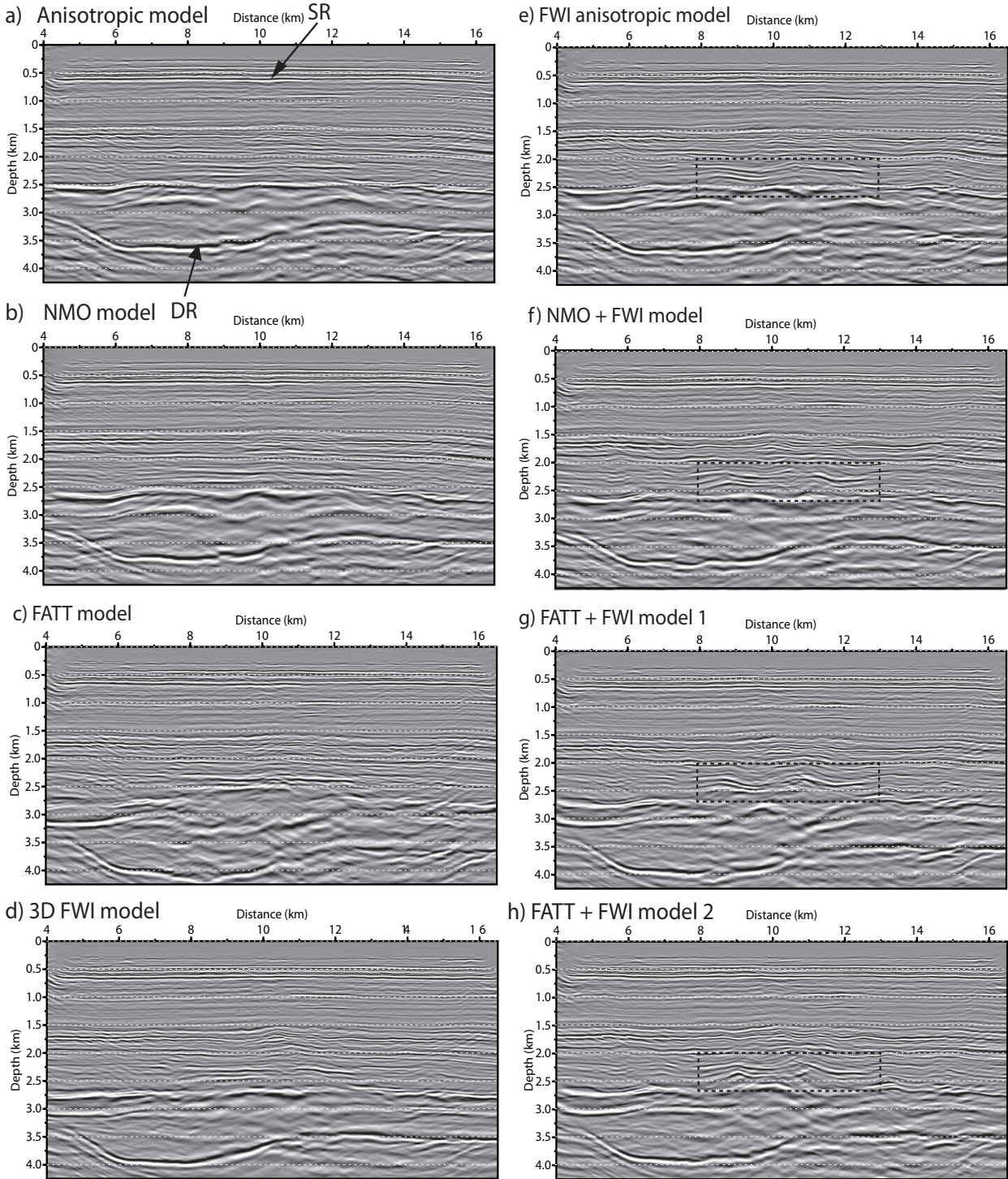


FIGURE 3.8 – Two-dimensional isotropic and anisotropic RTM images. (a-e) Anisotropic RTM images computed in the initial (a) and final (e) (V_{P0}, δ, ϵ) FWI models. (b,f) Isotropic RTM images computed in the NMO model (b) and $NMO + FWI$ model (f). (c-g,h) Isotropic RTM images computed in the $FATT$ model (c) and the final $FATT + FWI$ models 1 (g) and 2 (h). (d) Isotropic RTM image computed in the dip section of the 3-D FWI model of Sirgue *et al.* (2010). The black dashed frames in (e-h) show reflectors with suspicious horizontal undulations, that might be related to the footprint of anisotropy (see text for details).

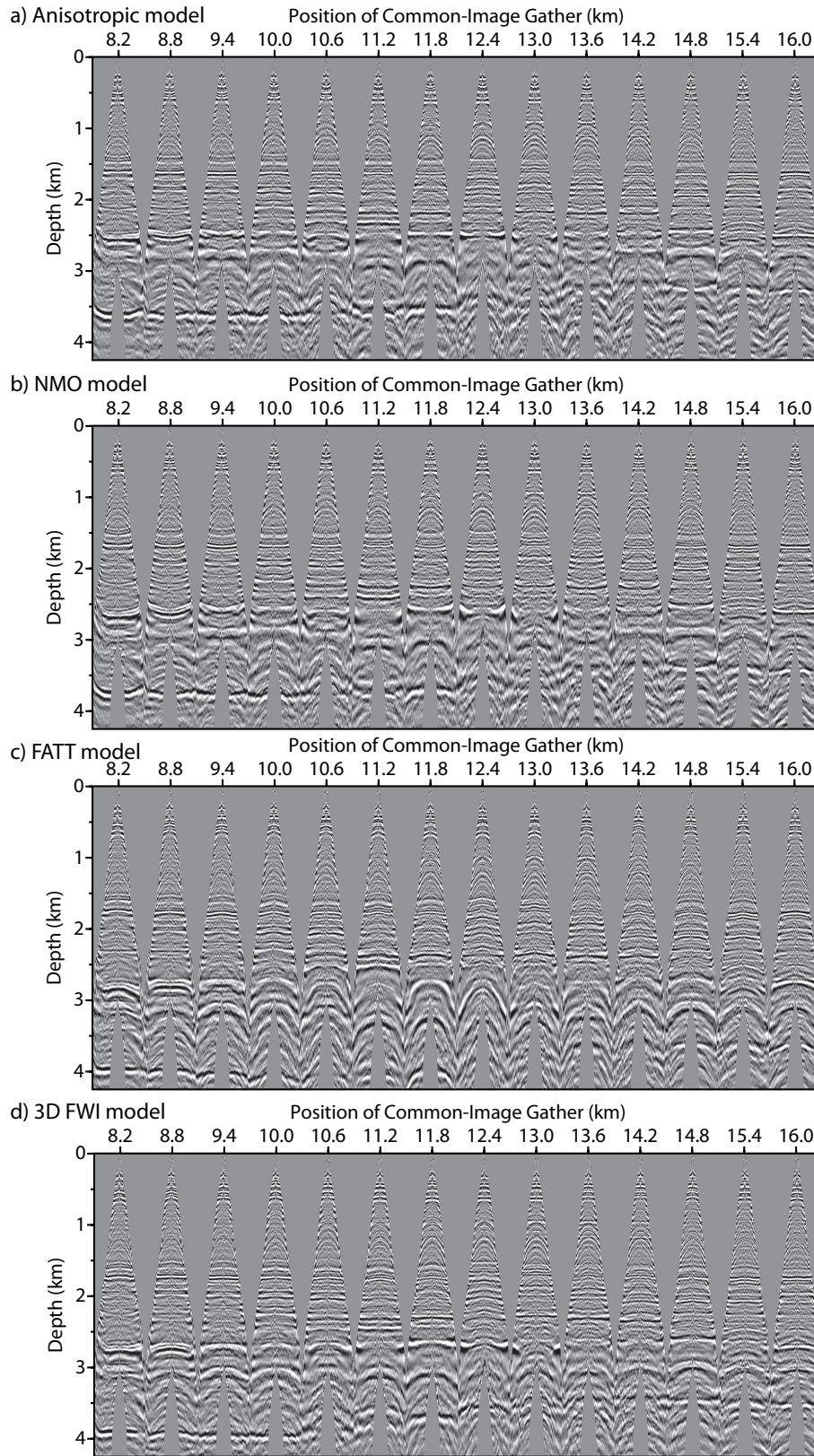


FIGURE 3.9 – CIGs corresponding to RTM images of Fig. 3.8. The horizontal axis gives the horizontal position of the CIG. Minimum and maximum offsets are -5 km and 5 km, respectively. Internal and external mutes were applied to the CIGs, which are plotted with an automatic gain control.

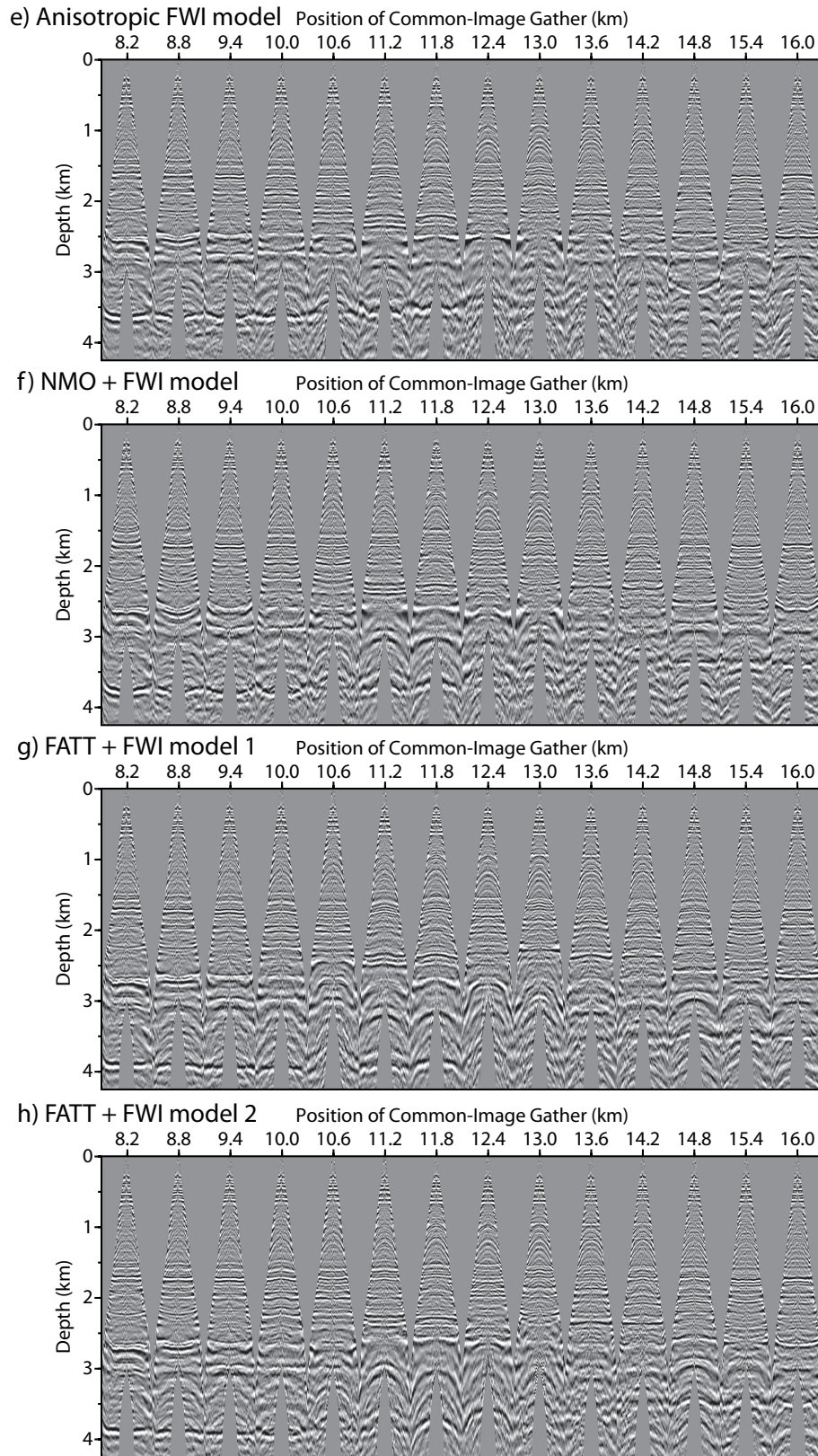


FIGURE 3.10 – Following of Fig. 3.9.

The migrated images computed in the FWI models are shown in Fig. 3.8(e-h). Overall, the anisotropic FWI model does not allow us to improve the RTM image of the deep structure obtained from the anisotropic reflection traveltime tomography (compare Fig. 3.8a and 3.8e). This is, to some extent, expected because the workflow, which combines the traveltime reflection tomography (or, migration-based velocity analysis) with the RTM, is more consistent than the one combining FWI with RTM. In the first case, the same subset of data, i.e., the short-spread reflections, is used during both the velocity model building and RTM, and the scale separation underlying these two tasks contributes to make the workflow well posed. In contrast, the FWI is a more ill-posed problem, where significant errors can be propagated in depth as longer offsets are processed. We note, however, that the reflectors in the CIGs inferred from the FWI model are significantly flatter in the shallow part (the first 1 *km*) than the ones inferred from the anisotropic reflection traveltime tomography model (Fig. 3.11). This highlights the capability of FWI for exploiting shallow reflections over the full aperture range, unlike reflection traveltime tomography. The improvement in the imaging of the shallow structure is observed for all of the migration tests described below. The *NMO + FWI* model produces CIGs, for which the smiling effects are slightly reduced (compare Fig. 3.9b and 3.9f). This is consistent with the increase in the velocities shown in the *NMO + FWI* model compared to the NMO model. Similarly, the *FATT + FWI* models help to reduce the frowning effects in the CIGs inferred from the FATT model (compare Fig. 3.9c and 3.9(g-h)). As expected, the RTM images inferred from the FATT and the *FATT + FWI* models are less well focused than the anisotropic and NMO counterparts (Fig. 3.8). This probably reflects the footprint of the initial FATT model, the horizontal velocities of which are not suitable for migration. However, the *FATT + FWI* model 2, compared to the FATT model, improves the imaging of the deep reflector that is raised by 300 *m* (compare Fig. 3.8c and 3.8(g-h)). Note also how the deep reflector in the RTM images inferred from the *FATT + FWI* models is positioned at the same depths than the one inferred from the 3D FWI model between 4 *km* and 11 *km* of distance (compare the RTM images in Fig. 3.8d,g,h and the corresponding velocity logs in Fig. 3.7(c-e)).

A striking feature of the RTM images inferred from the FWI models is suspicious horizontal fluctuations of the reflectors above the reservoir level between 2 *km* and 2.5 *km* in depth (Fig. 3.8(e-h)). This trend appears to gradually increase from the anisotropic RTM image to the *FATT + FWI* model 2 RTM image (from Fig. 3.8e to 3.8h). Although many factors might explain these fluctuations, one possible explanation would be related to the footprint of anisotropy. Indeed, the footprint of the anisotropy in the FWI results is expected to increase from the anisotropic FWI to the *FATT + FWI* application 2. For anisotropic FWI, the inaccuracies in the δ and ϵ background models might introduce some small bias into the $V_{P0} + FWI$ model through some trade-offs between vertical velocity and ϵ . On the other hand, the time dampings used in the *FATT + FWI* inversion (model 2) might contribute to strengthen artifacts associated with anisotropy because conflicting subdatasets sensitive to different kinds of velocities (ranging from vertical velocities to horizontal velocities through NMO velocities) are inverted hierarchically rather than simultaneously. Indeed, degrading data redundancy during successive intermediate stages of the FWI should give more freedom to the inversion to inject in the models artificial features to accommodate anisotropic effects. These horizontal fluctuations might be related with the lateral heterogeneity of the FWI models at the base of the gas layers and on top of the reservoir, that seems to increase from the $V_{P0} + FWI$ model to the *FATT + FWI* model 1 (Fig. 3.6).

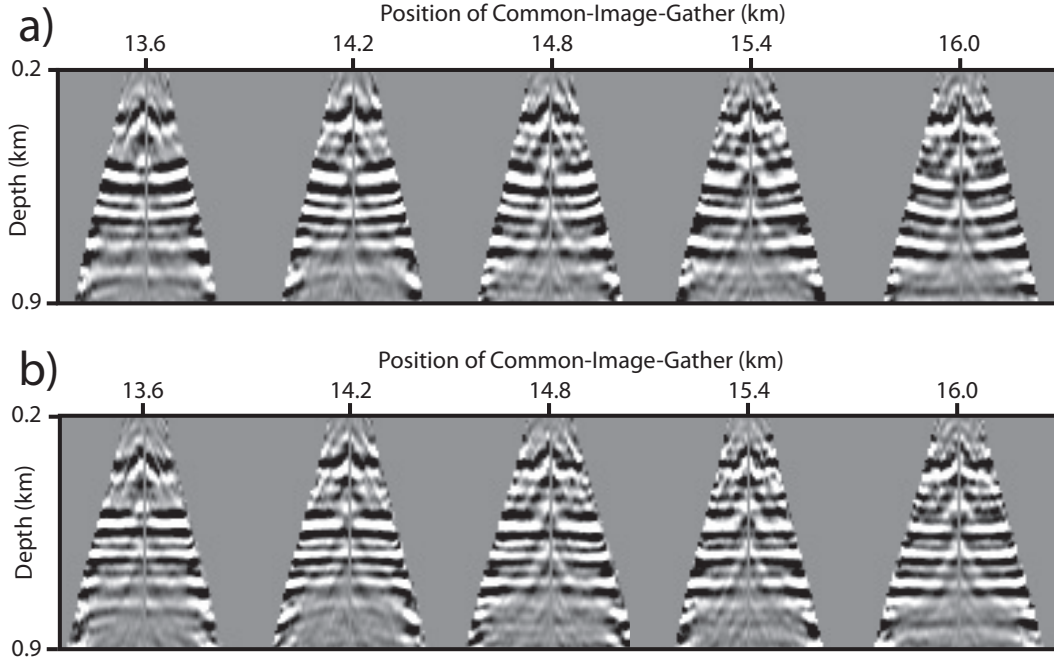


FIGURE 3.11 – Close-up of the CIGs of Fig. 3.9a (a) and Fig. 3.9e (b). The five most rightward CIGs are shown down to 900 m in depth. Note how the FWI allows for the flattening of shallow events.

Seismic modelling

For seismic modeling, we need a source wavelet, which is estimated by a VFSA method using a suitable subdataset and velocity model. We use the NMO velocity model and the first 2 km of offsets for the source estimation in the isotropic approximation. In this setting, the source-wavelet estimation is mainly controlled by the short-aperture reflection wavefield, which should be well predicted by the NMO model. The 220 sources wavelets associated with each receiver gather and the mean wavelet are shown in Fig. 3.12a. The source wavelet is computed within the 4-20-Hz frequency band. With the assumption of a uniform receiver-ground coupling all along the profile (a reasonable assumption according to the lithology of the sea bed, which is composed of hard sand (Kommedal *et al.*, 1997)) and a sufficiently-accurate velocity model, we should end up with quite similar source wavelets, as shown in Fig. 3.12a. In the following, we use the mean wavelet shown in the right panel of Fig. 3.12a for the seismic modeling.

Figure 3.13 shows a qualitative comparison between the receiver gather of Fig. 3.3a and synthetic seismograms computed in the final V_{P0} +FWI model (Fig. 3.6a) with the same DG modeling engine as for the FWI. Relatively good agreement is obtained between these two sets of seismograms, although we note underestimated amplitudes of the short-spread reflections Rg and Rs. This might be because, on one hand, it is generally acknowledged that data-space optimization by opposition to image-space optimization (or, migration-based velocity analysis) is not optimal to match short-spread reflections, (e.g., Symes, 2008), and, on the other hand, we consider only a smooth model for the density, which might have led to underestimate impedance contrasts.

A direct comparison between the recorded seismograms and the isotropic and anisotropic synthetic seismograms computed in the initial and final FWI models is shown in Fig. 3.14.

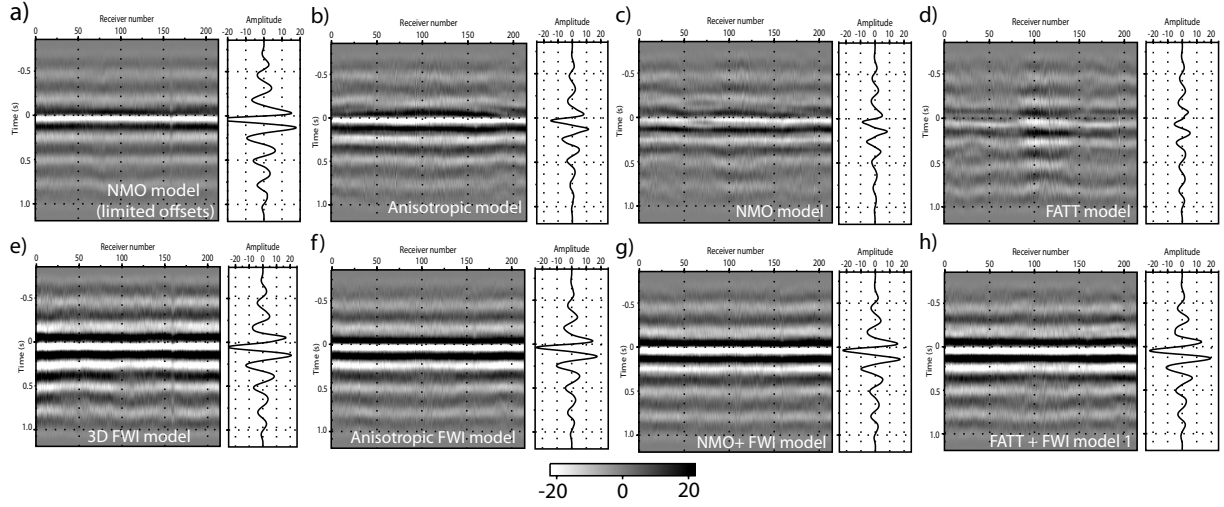


FIGURE 3.12 – Source-wavelets estimation. (a) Using isotropic modeling for a maximum offset of 2 km and the NMO velocity model. (b, f) Using anisotropic modeling for the full offset range and the initial (b) and final (f) (V_{P0}, δ, ϵ) FWI models. (c, g) Using isotropic modeling and the NMO model (c) and $NMO + FWI$ models (g). (d, h) As (c, g) for the FATT model (d) and the $FATT + FWI$ model 1 (h). (e) Using isotropic modeling for the full offset range and the dip section of the 3D FWI model of Sirgue *et al.* (2010). Note the improved focusing of the wavelets when using the FWI models.

The first obvious conclusion is that the data match obtained with the final FWI models (Fig. 3.14(e-h)) has been significantly improved compared to that obtained with the initial FWI models (Fig. 3.14(a-c)). A second important conclusion is that the data fits achieved with the $V_{P0} + FWI$ model, the $NMO + FWI$ model, and the $FATT + FWI$ models 1 and 2 are close, disregarding the discrepancies between these models. This thereby highlights the ill-posedness of the FWI in terms of non-unicity of the models to match anisotropic data. For all of the simulations performed in the FWI models, a reasonable match of the diving waves (D1, D2) and the reflections (Rs, Rg, Rr) has been achieved. The amplitudes of the diving waves (D1, D2) are, however, better matched in the seismograms computed in the $FATT + FWI$ models than in the seismograms computed in the $NMO + FWI$ model (compare Fig. 3.14f and 3.14(g-h)). This probably reflects the kinematic accuracy of the $FATT$ model to match first-arrival traveltimes. However, the seismograms computed in the $NMO + FWI$ model do not show obvious evidence of cycle skipping artifacts at large offsets for the diving waves, as the first-arrival traveltimes computed in the $NMO + FWI$ model accurately match the observed first-arrival traveltimes (Fig. 3.14f). The match of the first arrivals is consistent with our showing that the FWI converged towards velocities close to the horizontal velocities in the upper structure when the NMO model is used as the initial model (Fig. 3.7b). The successful match of the diving waves recorded at long offsets is unexpected when the NMO model is used as the initial FWI model because the traveltime mismatch between the recorded and the computed first-arrival traveltimes (0.3 s) exceeds the cycle-skipping limit at long offsets for a starting frequency of 3.5 Hz. This successful match can be interpreted on the basis that the FWI has performed a hierarchical layer-stripping reconstruction of the velocity structure over iterations, where the shallow part of the medium constrained by the high-amplitude short-offset early arrivals are reconstructed first followed by the reconstruction of the deeper part. This

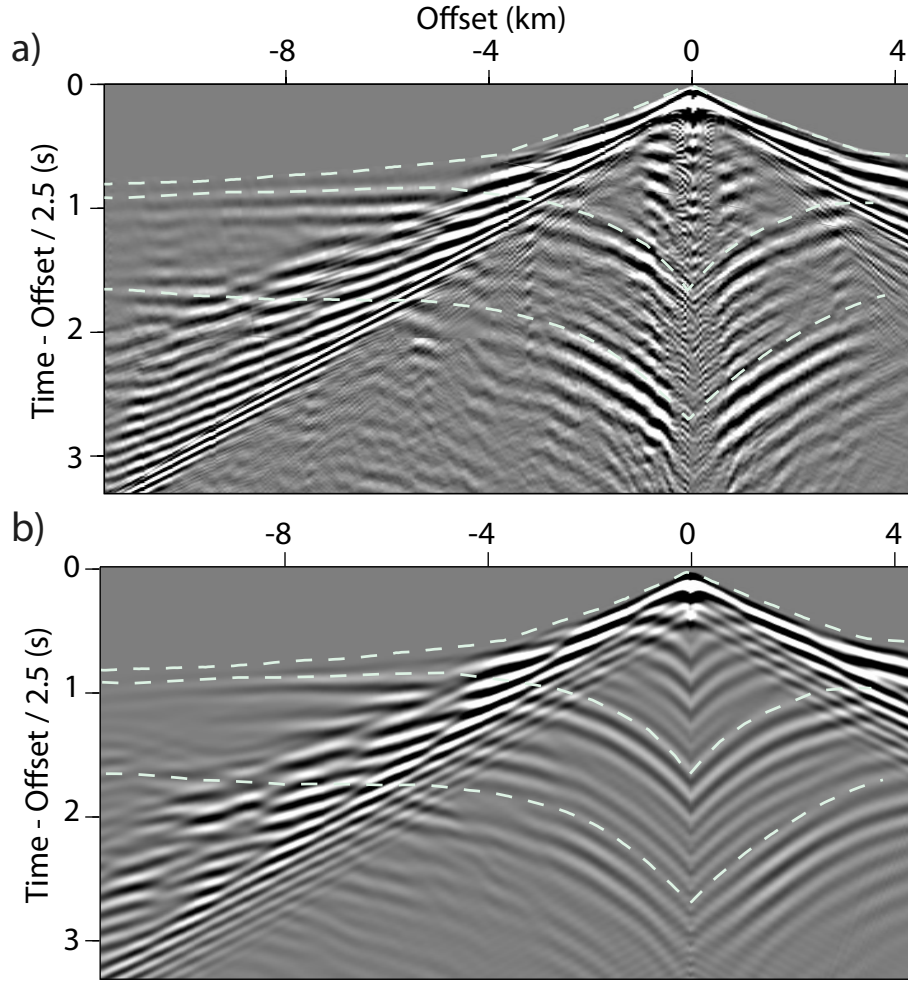


FIGURE 3.13 – Anisotropic modeling. (a) Recorded receiver gather of Fig. 3.3a. (b) Receiver gather computed in the $V_{P0} + FWI$ model. The dash lines show the first-arrival traveltimes computed in the FATT model and manually picked reflection traveltimes for phases R_g and R_r .

hierarchical reconstruction of the velocity structure according to depth might have contributed to the progressive absorption of the traveltime misfit with the offset, as the shallow part of the medium is improved by the FWI. The match of the short-spread reflection phases is almost equivalent in all of the seismograms, which is consistent with the concept that all of the initial models allow the prediction of the traveltimes within the cycle-skipping limit.

The data match is also shown by the plot of the misfit functions at the first and last FWI iterations as a function of the frequency group. The curves show essentially similar trends for the anisotropic FWI and the isotropic FWI using the NMO and FATT models as initial models (Fig. 3.15).

Source-wavelet estimation as a tool for model appraisal

We use the source-wavelet estimation as a tool to appraise the relevance of the FWI models

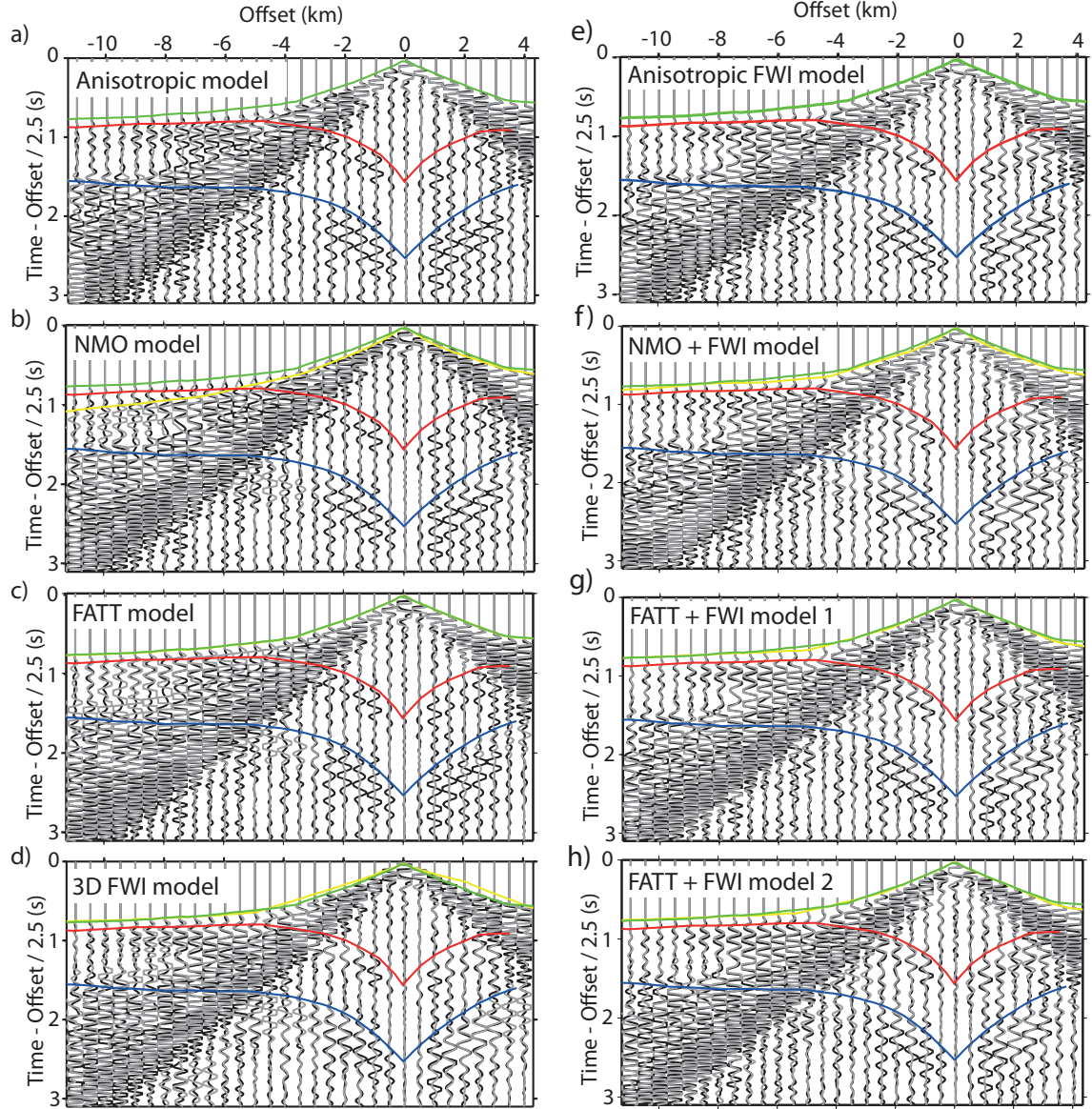


FIGURE 3.14 – Direct comparison between recorded (black) and modeled (gray) seismicograms for the receiver gather of Fig. 3.3a. (a, e) Anisotropic synthetic seismicograms computed in the initial (a) and final (e) (V_{P0} , δ , ϵ) FWI models. (b, f) Isotropic synthetic seismicograms computed in the NMO model (b) and in the $NMO + FWI$ model (f). (c, g, h) Isotropic synthetic seismicograms computed in the FATT model (b) and in the $FATT + FWI$ models 1 (g) and 2 (h). (d) Isotropic synthetic seismicograms computed in the dip section of the 3D FWI model of Sirgue *et al.* (2010). The green lines denote manually-picked first-arrival traveltimes. The yellow lines denote the computed first-arrival traveltimes in the current model. The red and blue lines are manually-picked reflection traveltimes for the phases Rg and Rr.

(Jaiswal *et al.*, 2009). We estimate the source wavelets by considering the full offset range, to make the wavelet estimation more sensitive to the model quality (Fig. 3.12). The sensitivity of the wavelet estimation to the amount of data used in Eq. 3.8 can be assessed by comparing

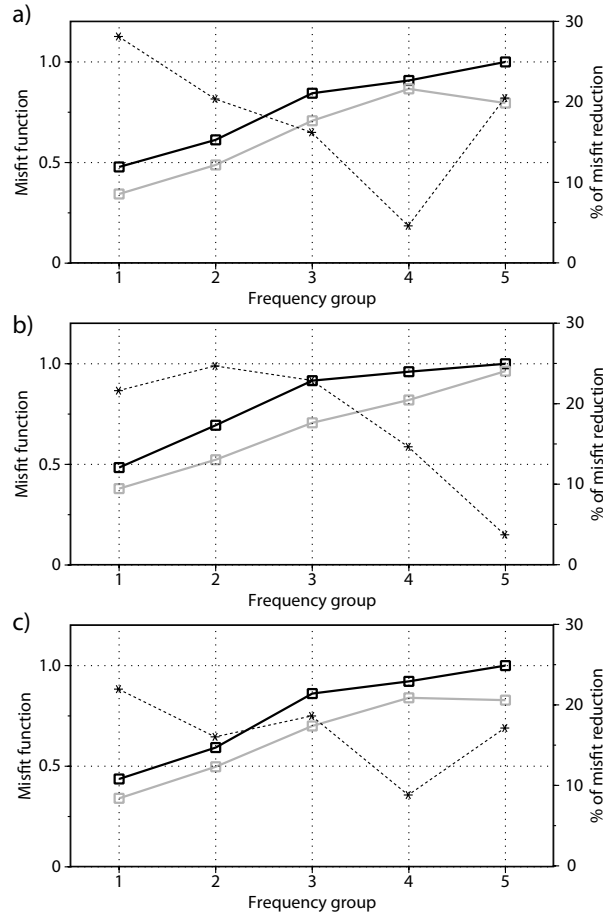


FIGURE 3.15 – Misfit function at first (solid gray) and last (solid black) iterations versus frequency groups for the anisotropic FWI for V_{P0} (a), the isotropic FWI using the NMO (b) and the FATT (test 3) (c) model as initial models. The dash line represents the percentage of misfit reduction.

the wavelets estimated in the NMO model using maximum offsets of 2000 m (Fig. 3.12a) and 13000 m (Fig. 3.12c). If the velocity model is not accurate enough, the repeatability of the wavelets is strongly affected when all of the aperture components of the data are involved in the inversion process. The collection of the source wavelets and the corresponding mean wavelet inferred from the initial and final FWI models when the full offset range is involved in the estimation is shown in Fig. 3.12(b-d), (f-h). Comparisons between the wavelets computed in the initial and the final FWI models show how the source-wavelet estimation is improved when a FWI model is used, hence validating the relevance of the FWI results. The wavelet inferred from the dip section of the 3D FWI model of Sirgue *et al.* (2010) is shown in Fig. 3.12e. This has a similar shape and amplitude to that obtained by 2D FWI.

3.2.5.6 Discussion

We have presented here the application of VTI anisotropic and isotropic acoustic FWI to wide-aperture OBC data from the Valhall field. We used both NMO and FATT models as initial models for the isotropic FWI, where the FATT model roughly represents the horizontal

velocities of the VTI medium. The results highlight the footprint of anisotropy on isotropic FWI.

Although there are significant differences between the kinematic properties of the initial models we used, the final isotropic FWI models show some overall common features. The most obvious one is that the horizontal velocities are mainly reconstructed in the upper structure, whatever the initial model and the data preconditioning. Therefore, the isotropic velocities in the upper structure are significantly higher than the velocities reconstructed by the mono-parameter anisotropic FWI for the vertical velocity. The horizontal velocities are also reconstructed in the 3D isotropic FWI model of Sirgue *et al.* (2010), and, therefore they cannot be interpreted as the footprint of the 3D effects. The reconstruction of the horizontal velocities in the upper structure shows that the FWI imaging is dominated by the diving waves and long-spread reflections, which sample the first 1.5 km of the medium. As horizontal velocities are reconstructed in the upper structure, this should impact upon the reconstruction of the deep structure that is mainly controlled by short-aperture reflections. Two pieces of evidence relating to this footprint might be low velocities in the gas layers (equal or even lower than the vertical velocities) and the stretching in depth of the velocity structure at the reservoir level. The stretching effect, which is clear in the 3D FWI results, is well illustrated in the RTM images where the depth of a deep reflector below the reservoir level gradually increases when we move from anisotropic migration to isotropic migration computed in the *FATT + FWI* models, these latter having the highest velocities in the upper structure. It is worth noting that the trend towards low velocities in the gas and deepening of the reservoir level tends to be strengthened when the prewhitening factor of the gradient preconditioning performed by the diagonal approximate Hessian is relaxed (see also Ravaut *et al.*, 2004, their Fig. 15) for the effects of the prewhitening factor on FWI). This suggests that the low velocities in the gas and the deepening of the top of the reservoir describe a real trend of the FWI convergence for this case study.

Although the 3D FWI model has been constructed from the NMO velocity model, the 3D FWI model shares more similarities with the 2D *FATT + FWI* models than with the *NMO + FWI* model. This is clearly shown by both the velocity-depth structure at the well log position (Fig. 3.7) and by the depths of the reflectors in the RTM images (Figs 3.8 and 3.9). This might arise because in 3D wide-azimuth acquisition, wide-aperture components have a stronger weight in the data than the short-aperture components, compared to 2D wide-aperture acquisitions.

The third conclusion is related to the ill-posedness of the FWI. We have shown that all of the FWI models provide almost equivalent matches of the data at short and wide apertures. Surprisingly, the FWI model inferred from the NMO initial model provides an acceptable match of the waveforms of the diving waves, despite the significant kinematic inaccuracy of the initial NMO velocities at wide apertures. The equivalent ability of the isotropic and anisotropic FWI models to match the data is also highlighted by the similarities of the source wavelet inferred from each FWI model. This raises the difficult issue of anisotropic acoustic FWI where several classes of anisotropic parameters should be reconstructed. Plessix & Cao (2011) and Gholami *et al.* (2011b) concluded that only two anisotropic parameters, related to the either the vertical or the NMO velocity and the horizontal velocity, can be reconstructed. Further demonstrations with realistic synthetic and real data examples need to be performed (Gholami *et al.*, 2011a). This will be the aim of future studies.

This study also illustrates that reflection data alone are not suitable for building horizontal

velocities in anisotropic media. We first show that the VTI model built by reflection tomography of short-spread reflection data does not match very accurately the first-arrival traveltimes in particular at short and intermediate offsets (Fig. 3.3b). The accuracy of the anisotropic velocity model in the shallow part is further improved by FWI, which contributes to flatten the CIGs in the shallow part of the medium (Fig. 3.11). Therefore, combining anisotropic refraction and reflection tomography should provide a more suitable framework to build an accurate initial model for FWI from wide-aperture data. Extension of anisotropic stereotomography to the joint inversion of refraction and reflection traveltimes is one possible approach to achieve this goal (Prioux *et al.*, 2010).

3.2.5.7 Conclusion

The case study of isotropic FWI of the anisotropic wide-aperture data presented in the present study has highlighted the footprint of anisotropy on isotropic FWI. We have used a surface real-data case study to discuss the validity of the isotropic approximation to perform FWI of wide-aperture/wide-azimuth data. For these acquisitions, when the data are modeled in the isotropic approximation, the differences between the horizontal and the vertical velocities leads to kinematic inconsistencies between the short-aperture and the wide-aperture components of the data. We show that the isotropic inversion is steered towards the reconstruction of the horizontal velocities in the upper part of the structure, which is sampled by diving waves and long-spread reflections. The reconstruction of the horizontal velocities in the upper structure leads to kinematic inconsistencies during the inversion of the short-spread reflections from deep reflectors, which are mainly sensitive to the vertical and the normal moveout velocities. These kinematic inconsistencies are accommodated by underestimated velocities and/or stretching in depth of the deep structure. Therefore, anisotropy should be involved in FWI to avoid this bias in the velocity estimation and in the depth positioning of the reflectors. Indeed, this is not a trivial task. We have shown how significantly different velocity models allow for a nearly equivalent match of the data. This highlights the ill-posedness of multiparameter acoustic anisotropic FWI. Therefore, future work will require a careful sensitivity analysis of the anisotropic FWI to define the number and type of parameter classes that can be reliably reconstructed by anisotropic FWI of wide-aperture data, as well as to design efficient strategies to constrain the inversion with suitable prior information coming from well logs.

Acknowledgements

This study was funded by the SEISCOPE consortium <http://seiscope.oca.eu>, sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, SAUDI ARAMCO, SHELL, STATOIL and TOTAL. The linear systems were solved with the MUMPS package, which is available on <http://graal.ens-lyon.fr/MUMPS/index.html>. The mesh generation was performed with the help of TRIANGLE, which is available on <http://www.cs.cmu.edu/~quake/triangle.html>. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and to the HPC resources of [CINES/IDRIS] under the allocation 2010- [project gao2280] made by GENCI (Grand Equipement National de Calcul Intensif). We gratefully acknowledge both of these Facilities and the support of their staff. We thank BP Norge and Hess Norge for providing us the 2D raw Valhall data set as well as the data set preprocessed by PGS for migration, the initial anisotropic models, the well log velocities

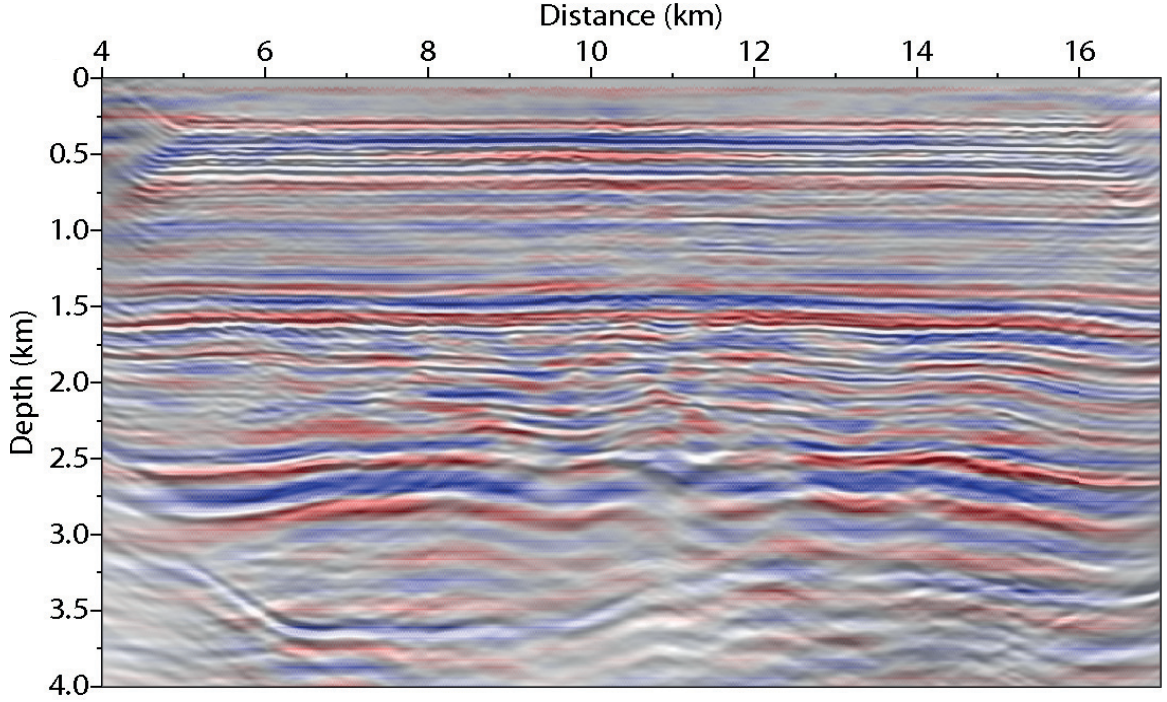


FIGURE 3.16 – Image migrée par RTM anisotrope dans le modèle final obtenu par inversion des formes d’ondes anisotropes, sur laquelle nous avons superposé le modèle de perturbation, défini comme la différence entre le modèle final et le modèle initial, et que nous avons ensuite dérivé verticalement pour augmenter la résolution verticale.

and the 2D section of the 3D FWI model of Sirgue *et al.* (2010). We would like to thank the associated Editor, Jeannot Trampert, an anonymous reviewer and R. -E. Plessix for their very constructive comments.

3.2.6 Résultats complémentaires à l’article

Pour compléter les illustrations de l’article, je montre, sur la Figure 3.16, la superposition de l’image migrée par RTM (anisotrope) calculée dans le modèle anisotrope d’inversion des formes d’ondes, et du modèle de perturbation de vitesses issu du même test. Nous constatons la bonne correspondance des structures visibles sur les deux images, tels que le toit de l’anticlinal ou bien le toit du réflecteur profond sous 3.5 km de profondeur. On notera le gap de résolution entre l’image migrée construite à partir du contenu haute fréquence des données et le modèle FWI construit à partir du contenu basse fréquence.

Sur les effets de l’anisotropie sur la loi de vitesse en fonction de la profondeur

Une manière de quantifier l’effet cinématique produit par une loi de vitesse $v(z)$, est de convertir les vitesses en temps de trajet en fonction de z : $t(z) = 2 \int_{z=0}^z (1/v(z)) dz$. Plus précisément, je vais analyser l’écart entre les temps double déduits du modèle \mathbf{m} et du modèle V_{NMO} : $\Delta t_{\mathbf{m}}(z) = t_{\mathbf{m}}(z) - t_{V_{NMO}}(z)$, où il est rappelé que le modèle V_{NMO} représente une vitesse intermédiaire entre la vitesse horizontale du modèle FATT et la vitesse verticale représentée par le log sonique. Nous montrons sur la Figure 3.17 les courbes correspondantes aux logs des

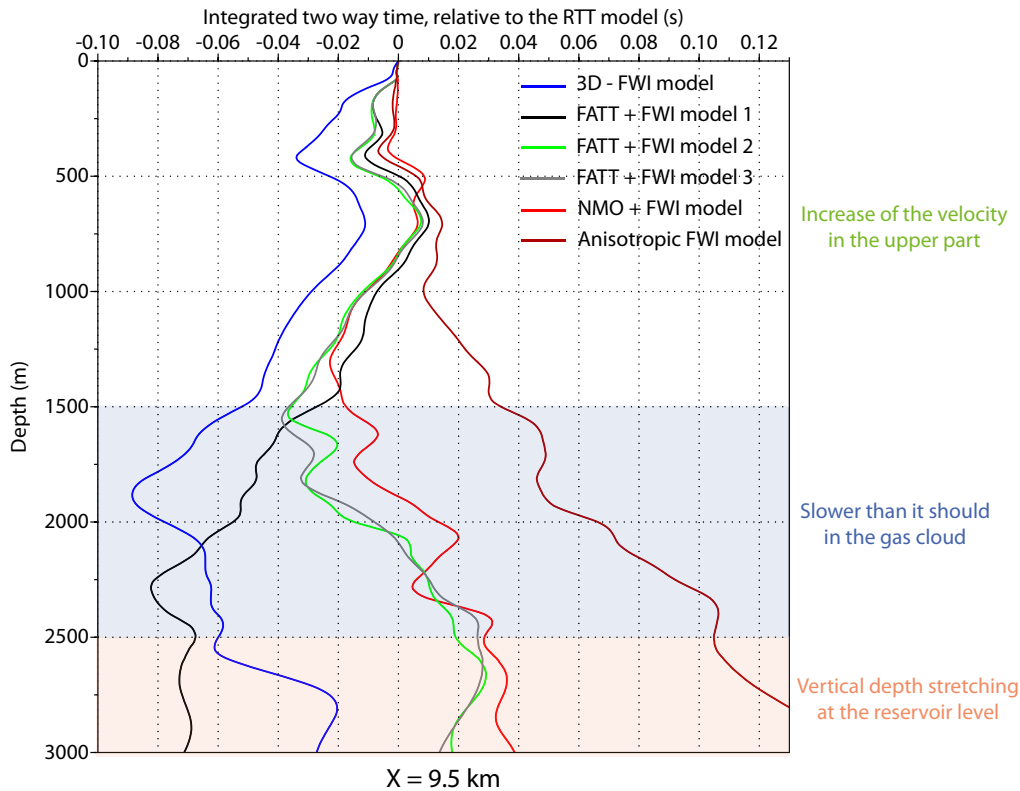


FIGURE 3.17 – Temps doubles intégrés le long des profils (à $X = 9.5\text{km}$) issus des différents modèles d'inversion des formes d'onde (Figure 3.7), auxquels furent soustraits les temps intégrés le long du profil de V_{NMO} . Les résidus positifs indiquent donc que le modèle considéré est plus lent que le modèle V_{NMO} , comme illustré par la courbe du modèle inversé anisotrope. Les variations des courbes des modèles isotropes montrent comment les effets de l'anisotropie sur la loi de vitesse évoluent en fonction de la profondeur.

différents modèles isotropes et du modèle anisotrope d'inversion des formes d'onde montrés dans l'article (Figures 3.6 et 3.7).

On constate que les résidus des temps pour tous les modèles isotropes diminuent et sont négatifs jusqu'à 1.5 km de profondeur, indiquant que ces modèles deviennent graduellement plus rapides que le modèle V_{NMO} avec la profondeur. Cette tendance est liée à l'augmentation de vitesse de la partie superficielle, décrite dans l'article comme une empreinte de l'anisotropie pour satisfaire l'accord avec les premières arrivées. Entre 1.5 et 2.5 km de profondeur, on remarque que les résidus associés aux modèles '*FATT + FWI model 2, 3*' et '*NMO + FWI model*' augmentent jusqu'à devenir positifs, ceci résultant des vitesses relativement faibles de ces modèles dans le nuage de gaz. Une inflexion positive en dessous de 2.5 km de profondeur, faible pour tous les modèles 2D isotropes, et très forte pour le modèle '*3D-FWI*', indique l'étirement vertical de la structure. Nous voyons ainsi comment, dans l'approximation isotrope, les vitesses rapides reconstruites dans la partie supérieure sont compensées dans la partie profonde par des vitesses anormalement faibles ou un étirement des structures en profondeur pour permettre l'ajustement des réflexions profondes. La différence significative des temps intégrés entre la courbe du modèle '*FATT + FWI model 1*' et celles des autres modèles isotropes, de l'ordre de

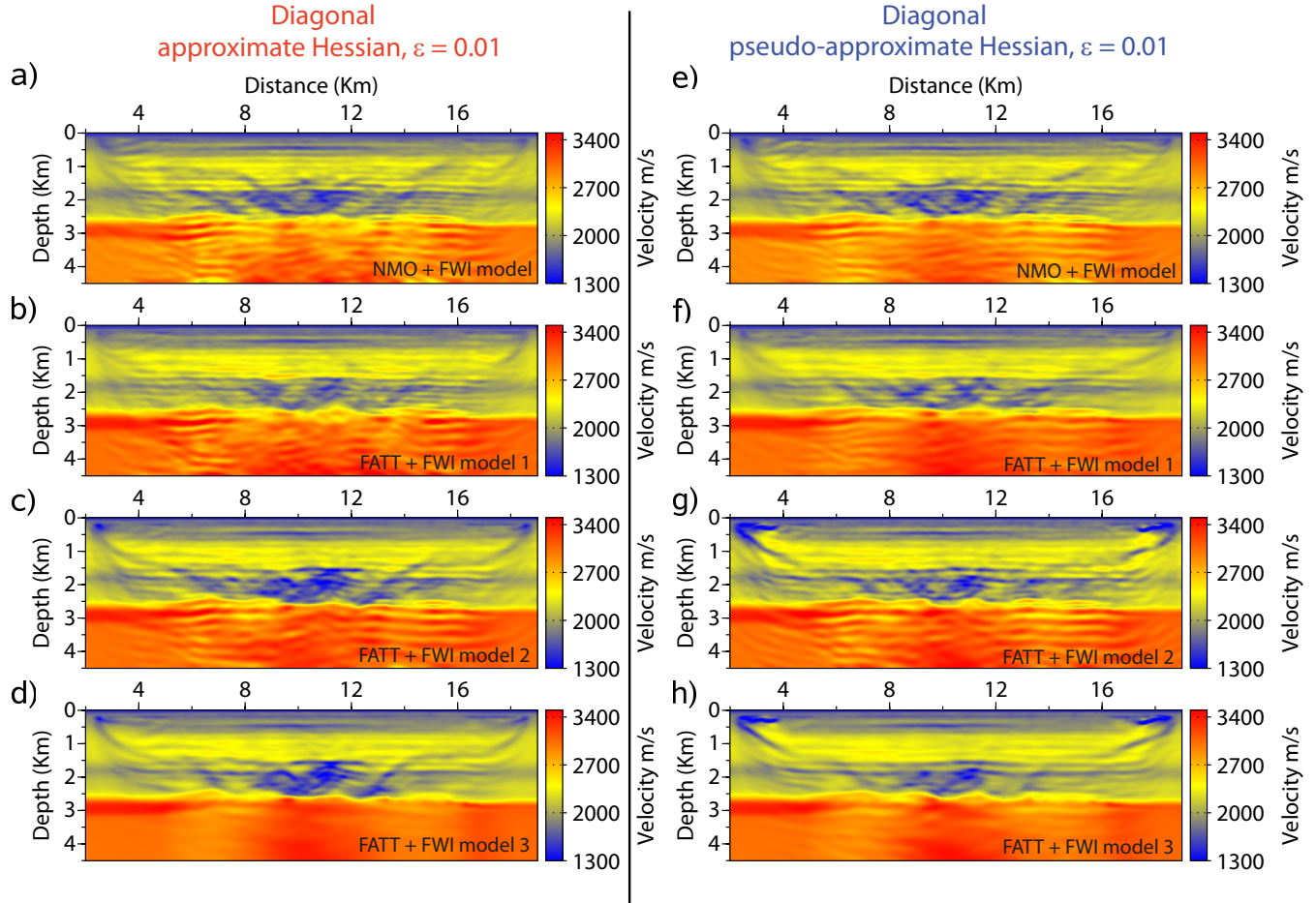


FIGURE 3.18 – (a-b) Modèles obtenus par inversion des formes d’ondes 2D isotrope en utilisant comme modèle initial le modèle V_{NMO} (a) et le modèle V_{FATT} (b-d) sans amortissement temporel (b) et avec trois (c) et un (d) amortissement temporel pour préconditionnement des données. (a-d) Tests obtenus en utilisant la diagonale du Hessien approximé, et (e-h) mêmes tests obtenus avec la diagonale du pseudo-Hessien approximé. Tous ont un facteur de préblanchiment égal à 0.01, contre 0.1 pour les tests de l’article, afin d’accentuer les perturbations.

0.1 s à 3 km de profondeur, n’est pas vraiment représentative de l’ensemble des modèles. Nous voyons en effet que le nuage de gaz est moins homogène que dans les autres modèles, et que le log passe à travers une zone de plus haute vitesse à cet endroit, ainsi qu’au travers d’une petite poche haute vitesse à 1.5 km de profondeur. L’étirement vertical du modèle est aussi plus visible aux positions $X = 9$ km et $X = 10$ km par exemple. Ce modèle reste néanmoins le plus rapide, comme l’atteste le calcul des sismogrammes calculés dans ce modèle, présentant un moins bon accord avec les données observées au niveau de la réflexion issue de la base du gaz (Figure 3.14g), probablement en raison de l’influence du modèle initial FATT utilisé.

Sur les effets des différentes approximations du Hessien

Je présente ci-après le résultat de tests de sensibilité à différents paramètres de réglage pour les inversions isotropes présentées dans l’article. Les tests de l’article ont été réalisés avec la diagonale du Hessien approximé (voir partie 1.2.3.3) et un facteur d’amortissement d’inversion

ϵ de 0.1 (Equation 1.62). Nous omettrons par brièveté dans cette partie le terme *diagonal* se référant aux approximations du Hessien considérées. Nous comparons ces tests (Figures 3.6 et 3.7) avec ceux obtenus en utilisant un facteur ϵ de 0.01 ajouté au Hessien approximé ou bien au pseudo-Hessien approximé (Shin *et al.*, 2001) (Figures 3.18 et 3.19). L'utilisation d'un facteur d'amortissement faible génère des perturbations plus fortes en profondeur, ce qui est conforme à l'effet attendu qui est de compenser l'expansion géométrique des fronts d'onde (Ravaut *et al.* (2004, leurs figures 12, 14, 15)). On peut par exemple suivre sur toute la largeur du modèle le réflecteur profond sous 3.5 km de profondeur dans le modèle '*NMO + FWI model*' obtenu avec le Hessien approximé (Figure 3.18a). Ces fortes perturbations dans la partie profonde accentuent les empreintes de l'anisotropie relevées dans l'article : les vitesses sont très faibles dans le gaz, en particulier dans les modèles '*FATT + FWI model 2*' et *3* obtenus avec le Hessien approximé ; et l'étirement de la structure en profondeur est plus marqué (Figure 3.19). Il est intéressant de constater que la diminution des vitesses dans le nuage de gaz est moins importante dans les modèles '*FATT + FWI model*' 2 et 3 obtenus avec le pseudo-Hessien, et que des artefacts basses vitesses sont apparus aux extrémités supérieures de ces modèles. Ces artefacts apparaissent aux endroits où l'acquisition n'est composée que de tirs, entre 2.5 km et 5 km de distance par exemple pour le bord gauche du modèle. Le pseudo-Hessien (Shin *et al.*, 2001) compense de manière moins précise l'expansion géométrique dans ces zones que le Hessien approximé, car le premier ne prend pas en compte les effets de propagation entre les récepteurs et les points du milieu (seulement entre les sources et les points). Ce degré de liberté laissé à l'inversion permet de diminuer artificiellement les vitesses aux extrémités supérieures du modèle, où un déficit d'éclairage existe, pour ajuster les réflexions à la base du gaz, ce qui a pour effet positif de reconstruire des vitesses dans le gaz plus réalistes que dans le cas où le Hessien approximé est utilisé. On notera également que ces artefacts n'apparaissent aussi fortement que pour les deux tests où sont utilisés les amortissements temporels. La raison en serait que l'inversion avec les amortissements temporels introduit des données aux informations cinématiques contradictoires de manière hiérarchique, et non simultanée. La décimation de la redondance des informations cinématiques laisserait alors plus de liberté à l'inversion pour créer des structures artificielles permettant d'accommoder les effets d'anisotropie. Ces artefacts semblent être corrélés à l'ondulation horizontale du toit de l'anticlinal, car les ondulations sont d'autant plus marquées que les artefacts sont importants.

3.2.7 Conclusion de l'article

Nous avons vu qu'il était primordial de comprendre la géométrie de propagation des différents types d'ondes se propageant dans le milieu, et leur signature dans les données. En effet, la distinction entre les ondes réfractées (et réfléchies super-critiques) et les ondes réfléchies pré-critiques, enregistrées respectivement à grandes et à faibles ouvertures, m'a conduit à compartimenter verticalement la loi de vitesse isotrope reconstruite dans le milieu. Des vitesses rapides, proches des vitesses horizontales, sont reconstruites dans la partie superficielle du milieu, où l'inversion est dominée par la contribution des arrivées grands angles qui se propagent de manière subhorizontale dans le milieu. Des vitesses anormalement faibles dans la structure profonde du milieu ou son étirement vertical sont alors nécessaires pour accommoder les vitesses superficielles rapides et ajuster les réflexions profondes, caractérisées par des ouvertures plus faibles en raison de l'extension limitée du dispositif d'acquisition. Ce jeu de vase communiquant rappelle la non-linéarité intrinsèque des méthodes d'imagerie en terme de non-unicité de la solution. Par ailleurs, cette étude souligne l'importance de prendre en compte l'anisotropie dans les

3.2 Empreinte de l'anisotropie du milieu sur une inversion acoustique isotrope

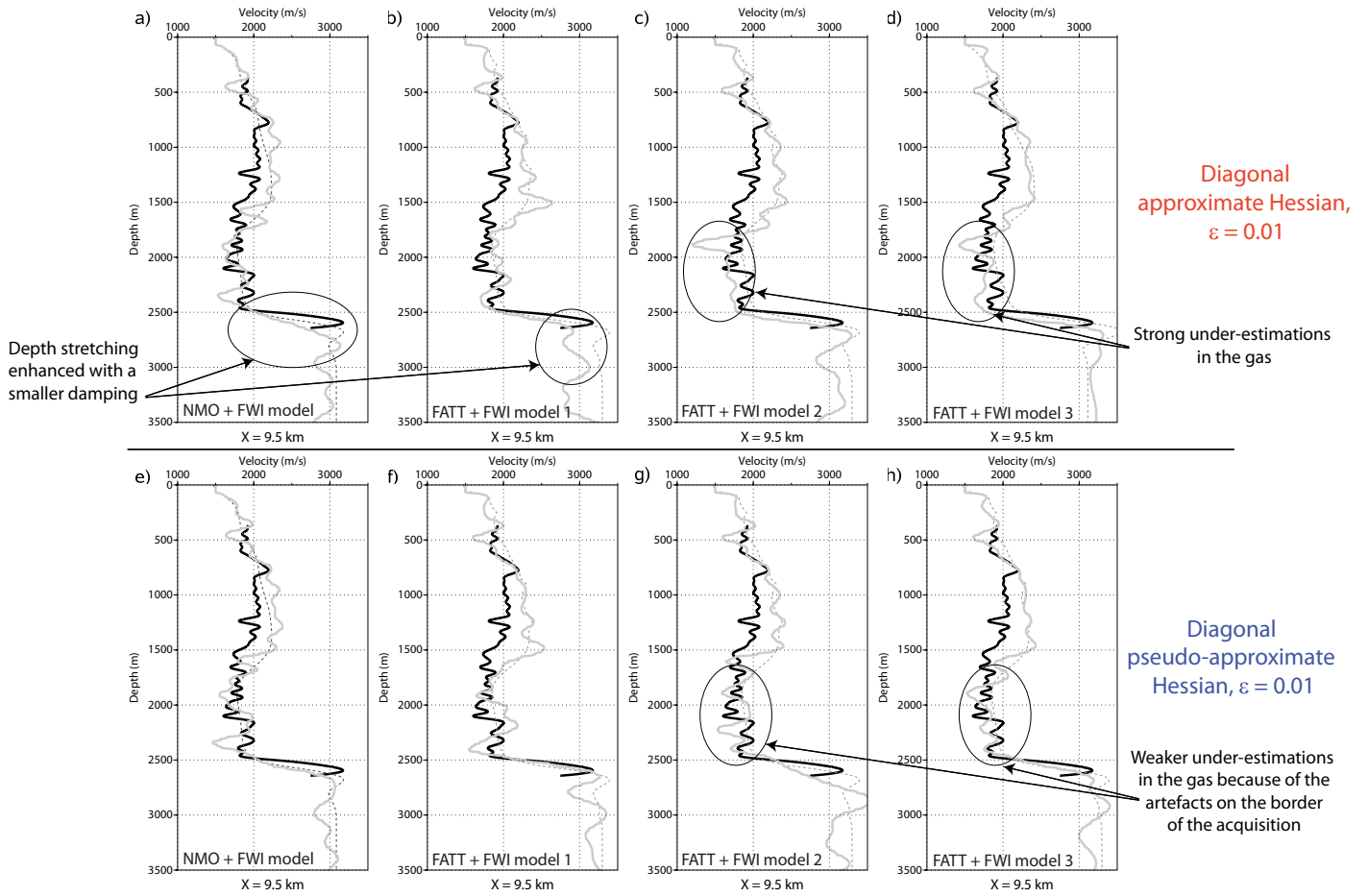


FIGURE 3.19 – (a-h) Profils extraits des modèles présentés sur la Figure 3.18(a-h).

modélisations, le biais des vitesses isotropes reconstruites ayant pour conséquence un mauvais positionnement des réflecteurs en profondeur. Ceci peut avoir des conséquences dramatiques pour la prospection pétrolière lors du forage de puits, bien qu'un calage aux puits permette en général de corriger ce biais dans les vitesses. Notons que la construction des modèles initiaux anisotropes, nécessaire à l'inversion des formes d'onde anisotrope, n'est pas toujours aisée, d'autant que des logs de puits sont potentiellement nécessaires pour calibrer la profondeur des réflecteurs. Cette étude sur l'empreinte de l'anisotropie d'une inversion isotrope est cas dépendante, bien que les milieux tabulaires comportant une anisotropie VTI constituent un cadre relativement réaliste. L'étude aura montré que l'extension à d'autres cas d'études passera par la compréhension du type d'anisotropie, de l'acquisition et des données correspondantes.

3.3 Analyse paramétrique de l'inversion acoustique

3.3.1 Non-linéarités introduites par la surface libre - apport du préconditionnement - sensibilité à l'offset

Cette partie constitue un des aspects traités dans un acte présenté au congrès de la SEG en 2009 (Prioux *et al.*, 2009). La modélisation utilisée dans cette partie est basée sur une méthode de différences finies dans le domaine des fréquences, adaptée aux milieux 2D visco-acoustiques isotropes (Hustedt *et al.*, 2004). Dans les tests suivants, neuf fréquences entre 2 et 15 Hz sont successivement inversées. L'ondelette source est un ricker centré sur la fréquence 7 Hz et est supposée connue. Le modèle initial fut construit par lissage gaussien avec une longueur de corrélation constante de 500 m. Dans un premier temps, nous n'utilisons pas d'amortissement temporel. Les multiples associés à la surface libre complexifient le champ d'onde, en particulier en présence d'une faible profondeur d'eau et lorsque de grandes distances de propagation sont considérées. En effet de nombreuses réflexions ont alors lieu entre la surface libre et les réflecteurs du fond marin et du sous-sol avant d'atteindre le récepteur. Ces réflecteurs n'étant pas forcément présents dans le modèle initial, il est difficile pour l'inversion de séparer la contribution des différents points où se sont produites les réflexions, d'autant que nous ne prenons en compte ici que la diagonale du premier terme du Hessien de l'équation 1.26, et que nous ignorons le deuxième terme du Hessien qui rend compte des phénomènes de diffractions doubles (Pratt *et al.* (1998) et Métivier *et al.* (2012a)). Il est intéressant de voir en quoi la présence de multiples dans les données va affecter la convergence de l'inversion. Pour cela nous calculons dans le vrai modèle deux jeux de données synthétiques caractérisés par un offset maximum de 32 km, l'un avec une condition absorbante et l'autre avec une surface libre au toit du modèle, que nous inversons en partant d'un modèle initial lisse. Précisons que le modèle de 32 km de long a été obtenu à partir du modèle synthétique original de 16 km de long, en étendant celui-ci latéralement de 8 km de part et d'autre. Nous observons sur les modèles finaux (Figure 3.20(a-b)) comment la non-linéarité introduite par la prise en compte de la surface libre s'est traduite par la création d'artefacts.

Pour réduire la non-linéarité de l'inversion des données à grands déports en présence de multiples, j'ai refait le test en utilisant quatre amortissements temporels appliqués à partir de la première arrivée, de valeurs $\tau = 0.5, 1, 2$ et 5 s, de manière à introduire progressivement les angles d'ouverture plus faibles dans l'inversion. Nous montrons la partie centrale du résultat sur la Figure 3.20c. Nous constatons que la reconstruction hiérarchique des nombres d'ondes a permis l'élimination des artefacts. Nous voyons sur la Figure 3.20d qu'une autre manière de faire disparaître ces artefacts liés à la surface libre est de ne considérer que 16 km d'offset. L'explication, originellement donnée par Sirgue (2003, 2006), vient du fait que la non-linéarité de l'inversion augmente avec le nombre de longueurs d'ondes propagées. Cela explique d'ailleurs probablement pourquoi les artefacts sont localisés dans les 2.5 premiers km de profondeur, dans lesquels se propagent les arrivées réfractées à grands offsets. Cependant, il faut rappeler que l'allongement du dispositif d'acquisition permet d'enregistrer de plus grandes ouvertures angulaires, propices à l'imagerie des grandes longueurs d'ondes du modèle. Pour illustrer ce point, j'ai comparé le résultat des tests obtenus avec condition absorbante, en utilisant 16 km et 32 km d'offset (Figure 3.21). Le profil extrait à la distance de 1.25 km montre que le modèle avec 16 km d'offset est très imprécis à cette position, en raison du déficit d'éclairage au voisinage du bord du modèle. Ce type d'artefact est aussi visible aux extrémités du modèle de la Figure 3.20d, issu du test obtenu avec 16 km d'offset et une surface libre. Les profils extraits

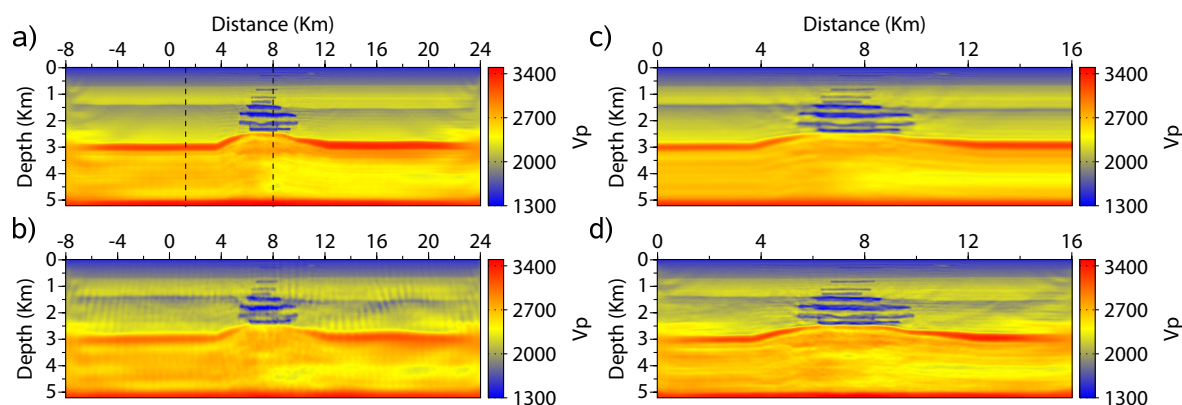


FIGURE 3.20 – Effets des multiples sur l'inversion des formes d'onde. (a-b) Modèles obtenus en utilisant 32 km d'offset lorsqu'une condition absorbante (a) et une condition de surface libre (b) sont implémentées au toit du modèle. (c,d) Modèles issus d'une inversion avec une surface libre, où sont utilisés 32 km d'offset et 4 amortissements temporels en (c) (seul la partie centrale est représentée), alors que sont utilisés 16 km d'offset sans amortissements en (d).

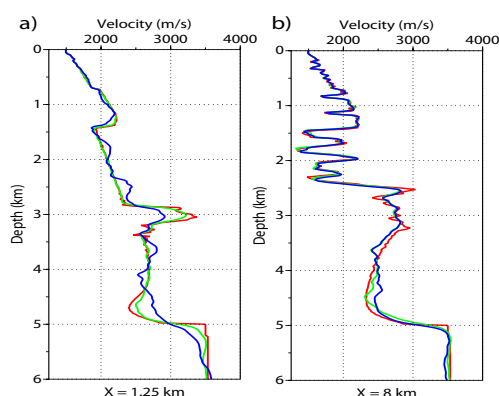


FIGURE 3.21 – Profils extraits à 1.25 km (a) et 8 km (b) de distance, du modèle synthétique de Valhall (courbe rouge), et des modèles finaux d'inversion des formes d'onde obtenus avec une condition absorbante, lorsque 16 km d'offset (courbe bleue) et 32 km d'offset (courbe verte) sont utilisés. Notez l'imprécision du modèle avec 16 km d'offset en raison des effets de bord de l'acquisition, et l'intérêt des grands nombres d'onde fourni par les dépôts de 32 km, pour reconstruire la partie centrale profonde du modèle.

au centre des modèles confirment une reconstruction partielle des grandes longueurs d'onde dans la partie profonde du modèle de 16 km, en raison d'un éclairage angulaire plus étroit.

3.3.2 Sensibilité aux modèles initiaux anisotropes, à la fréquence initiale et au préconditionnement des données, relativement à une inversion isotrope

L'analyse de sensibilité ci-après a pour objectif d'évaluer l'impact de la fréquence initiale (2 Hz ou 3.5 Hz) sur la reconstruction du modèle de vitesse synthétique de Valhall par inversion

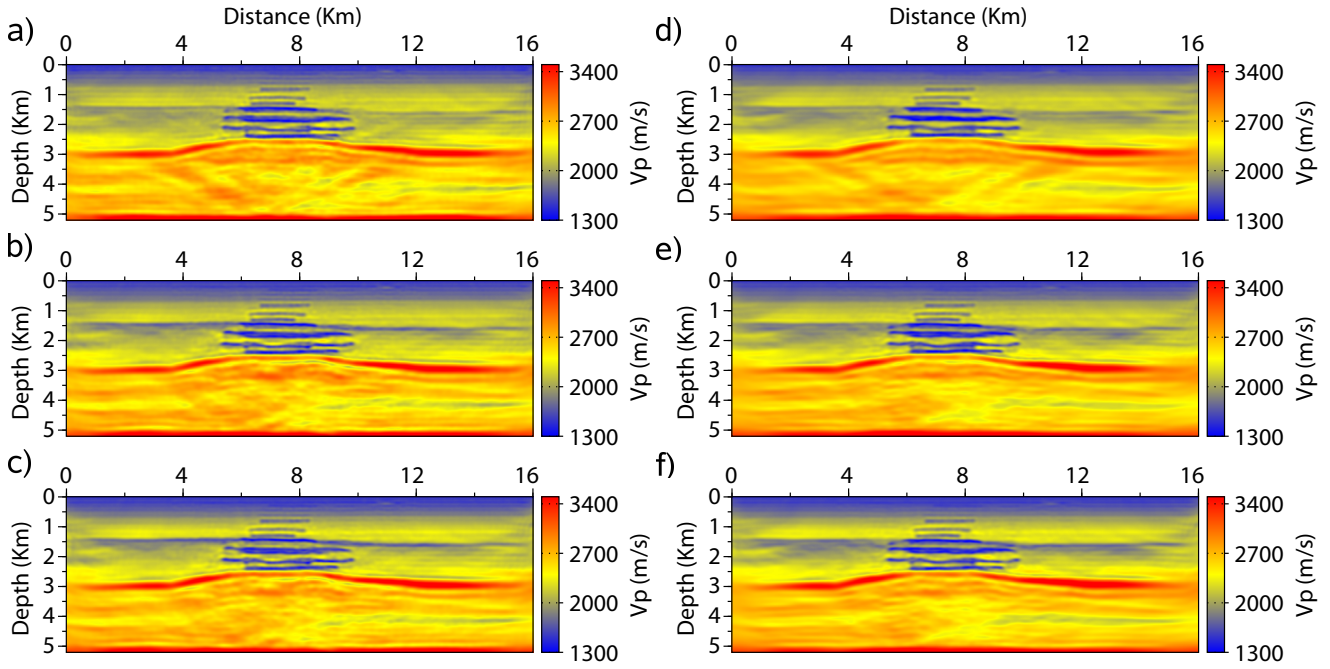


FIGURE 3.22 – (a,d) Modèle de V_P obtenu à l'issue d'une inversion utilisant une modélisation isotrope (cas 1) sans (a) et avec 3 amortissements de 1, 3 et 10 s (d). (b,e) Inversion anisotrope en utilisant comme modèle de référence des paramètres de Thomsen les modèles exacts (cas 2) sans (b) et avec 3 amortissements (e). (c,f) Idem que (b,e) mais pour des modèles de référence des paramètres de Thomsen obtenus par lissage des modèles exacts (cas 3). La fréquence initiale de ces tests est 2 Hz.

des formes d'onde, dans le cas des approximations isotrope (cas 1) et anisotrope (cas 2 et 3). Par la suite, je testerai également l'apport potentiel des amortissements temporels des données (τ) sur l'inversion. Pour les tests d'inversion anisotrope, les modèles des paramètres de Thomsen sont toujours maintenus fixes au cours de l'inversion, mais je distingue le cas où ces modèles de référence sont égaux aux modèles exacts (cas 2), et le cas où ils sont obtenus en lissant les modèles exacts (cas 3). Le modèle initial de V_{P0} , ainsi que les paramètres de Thomsen dans le cas 3, ont été générés en lissant le modèle exact avec des longueurs de corrélation horizontale et verticale de 500 m. Les fréquences inversées sont : [2, 3, 4, 5, 6.1, 7] Hz lorsque la fréquence initiale est égale à 2 Hz, et [3.5, 4, 5, 6.1, 7] Hz lorsque la fréquence initiale est égale à 3.5 Hz. Les amortissements temporels lorsqu'ils sont utilisés ont pour valeur 1, 3 et 10 s. La régularisation employée consiste à lisser le gradient de la fonction coût à chaque itération. J'utilise la diagonale du Hessien approximé pour le préconditionnement du L-BFGS. Les sources et les récepteurs sont espacés de 50 m, et sont localisés en surface à respectivement 6 m et 71 m de profondeur. Le nombre d'itération est de 30 pour les tests sans amortissement, et de 10 par amortissement pour les tests avec trois amortissements, afin de pouvoir comparer de manière objective ces deux types de tests.

Les modèles et leurs logs obtenus par inversion des formes d'onde en partant de 2 Hz avec et sans amortissement sont présentés respectivement sur les Figures 3.22 et 3.23. De manière générale on observe que les modèles obtenus sans amortissement (colonne de gauche)

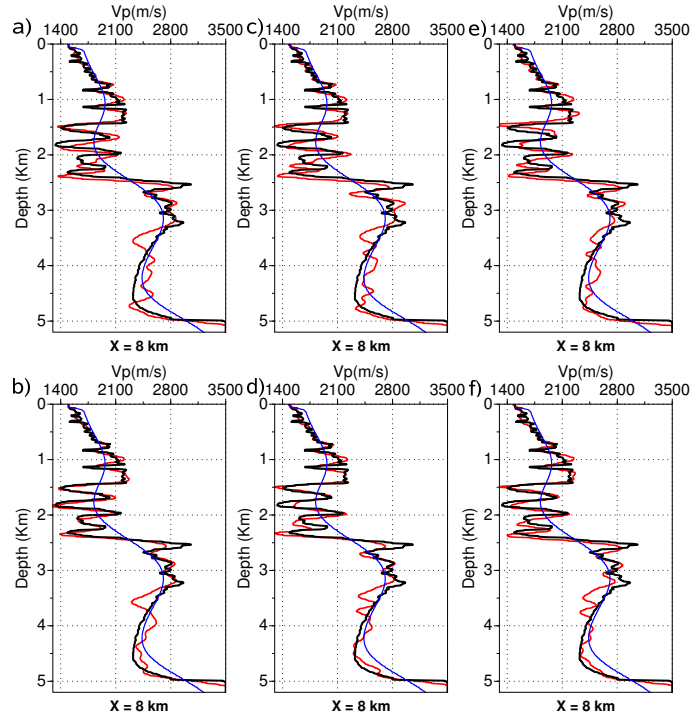


FIGURE 3.23 – (a,c,e) Logs extraits au centre des modèles issus des inversions sans amortissement dans les cas 1, 2 et 3 respectivement, montrés sur la Figure 3.22(a-c). (b,d,f) Idem pour les logs provenant des inversions avec amortissements. La différence entre les résultats avec et sans amortissement est de moins en moins importante du cas 1 au cas 3. Les courbes noires, rouges et bleues correspondent respectivement au vrai modèle, au modèle retrouvé par inversion des formes d’ondes et au modèle initial.

paraissent un peu moins lisses que les modèles obtenus avec amortissements. L’examen des logs nous indique que les modèles avec amortissements semblent être plus proches du vrai modèle que ceux sans amortissements, mais il apparaît aussi que cette différence est beaucoup plus prononcée dans le cas isotrope (cas 1) que dans le cas anisotrope où les paramètres de Thomsen sont lissés (cas 3). Pour avoir une estimation plus quantifiée de la qualité des modèles reconstruits, j’introduis le pourcentage d’erreur relatif ξ défini comme

$$\xi(\%) = \frac{100}{n_x \cdot n_z} \sum_{i=1}^{n_x \cdot n_z} \frac{|m_{cal,i} - m_{vrai,i}|}{m_{vrai,i}} \quad (3.9)$$

où $|\cdot|$ est la valeur absolue ; n_x et n_z le nombre de points suivant la direction x et z ; $m_{cal,i}$ et $m_{vrai,i}$ le modèle retrouvé par inversion et le vrai modèle au point i considéré. Nous allons maintenant analyser la table 3.2 rassemblant les pourcentages d’erreur de l’ensemble des tests. Notons que ces pourcentages d’erreur ont une signification réelle lorsque nous nous en servons pour juger la qualité d’un modèle par rapport à un autre, car tous les modèles ont une résolution comparable.

Une première conclusion attendue est que l’utilisation d’une fréquence initiale basse permet d’avoir une meilleure convergence du problème inverse. L’erreur est par exemple 1.77 % plus

faible en partant de 2 Hz que de 3.5 Hz pour les tests du cas 3 (avec les paramètres de Thomsen lissés). Cette différence d'erreur se trouve être plus faible lorsque des amortissements sont utilisés (environ 0.96%), ce qui illustre l'apport du préconditionnement hiérarchique des données par amortissement temporel pour réduire les non-linéarités du problème inverse. Par ailleurs, les tests anisotropes où les paramètres de Thomsen sont exacts sont toujours significativement moins précis que les tests isotropes, d'environ 0.33 % en moyenne sur tous les tests. Une raison possible est que la matrice d'impédance anisotrope est moins bien conditionnée, ce qui provoquerait de plus fortes erreurs numériques lors de la factorisation LU. Notons cependant que toutes les simulations anisotropes (acoustiques) sont réalisées en double précision, alors que les tests isotropes sont en simple précision. Par ailleurs, les PML peuvent être une source d'instabilités numérique dans les milieux anisotropes, en particulier pour les incidences rasantes (Meza-Fajardo & Papageorgiou, 2008). Une autre raison possible est liée aux non-linéarités introduites par les ondes S (parasites) converties aux interfaces dans le cadre de l'approximation acoustique anisotrope (Operto *et al.*, 2009).

Ces tests mettent aussi en évidence l'effet de l'imprécision des paramètres de Thomsen sur la qualité du modèle retrouvé. Afin de se rendre compte de l'effet du lissage des paramètres anisotropes sur les temps d'arrivée et les amplitudes, je présente sur la Figure 3.24 les sismogrammes calculés dans les modèles exacts de V_{P0} , de densité et d'atténuation, avec dans un cas les paramètres exacts de Thomsen et dans l'autre cas leur version lissée, ainsi que la différence entre ces sismogrammes. Cette différence montre nettement que l'imprécision des paramètres de Thomsen affecte principalement les données enregistrées à grands déports, et peu ou pas les réflexions. L'erreur cinématique liée au lissage des paramètres de Thomsen augmente donc avec le nombre de longueurs d'ondes propagées. Nous avons aussi superposé les hodochrones des temps de premières arrivées pointées manuellement sur les deux jeux de sismogrammes. On observe une différence de temps de première arrivée vers 12 km d'offset d'environ 0.1 s, ce qui est suffisant pour créer un saut de phase pour une fréquence initiale de 5 Hz. La conséquence de ces imprécisions se traduit par un pourcentage d'erreur 0.81 % plus fort pour le test du cas 3 que pour le test du cas 2, lorsque la fréquence initiale est 3.5 Hz et que l'on utilise des amortissements. Cette sensibilité aux erreurs cinématiques du modèle initial s'avère être moins marquée à 2 Hz qu'à 3.5 Hz ($\Delta\xi_{2Hz} = 0.3$ % contre $\Delta\xi_{3.5Hz} = 0.75$ % pour les tests sans amortissement). Par ailleurs, les tests avec amortissements sont plus sensibles au lissage des paramètres de Thomsen que les tests sans amortissement, comme l'atteste à 3.5 Hz, une différence de pourcentage d'erreur de 0.6 % avec amortissements contre 0.3 % sans amortissement. Bien que les tests avec amortissements présentent toujours une erreur plus faible que les tests sans amortissement, la stratégie des amortissements temporels est moins efficace si les paramètres de Thomsen ne sont pas exacts, car alors l'inversion a plus de degrés de liberté pour créer des structures artificielles pour expliquer les données à grands puis à courts angles. Ce n'est pas le cas si toutes les données sont introduites simultanément dans l'inversion (cas sans amortissement), car alors le processus d'optimisation explique 'au mieux' l'ensemble des données. Cela se traduit aussi par une différence entre les résultats avec et sans amortissement qui est moins grande dans le cas 3 que dans les cas 1 et 2, avec par exemple une différence d'erreur de 0.16 % dans le cas 3, contre 0.55 % et 0.46 % dans les cas 1 et 2, lorsque la fréquence initiale est de 2 Hz.

Nous avons pu quantifier l'influence de la fréquence initiale, du préconditionnement et de la précision des modèles initiaux anisotropes sur la qualité de l'inversion. La manière et la force avec laquelle les différents paramètres influent sur l'inversion sont bien sûr dépendantes du cas

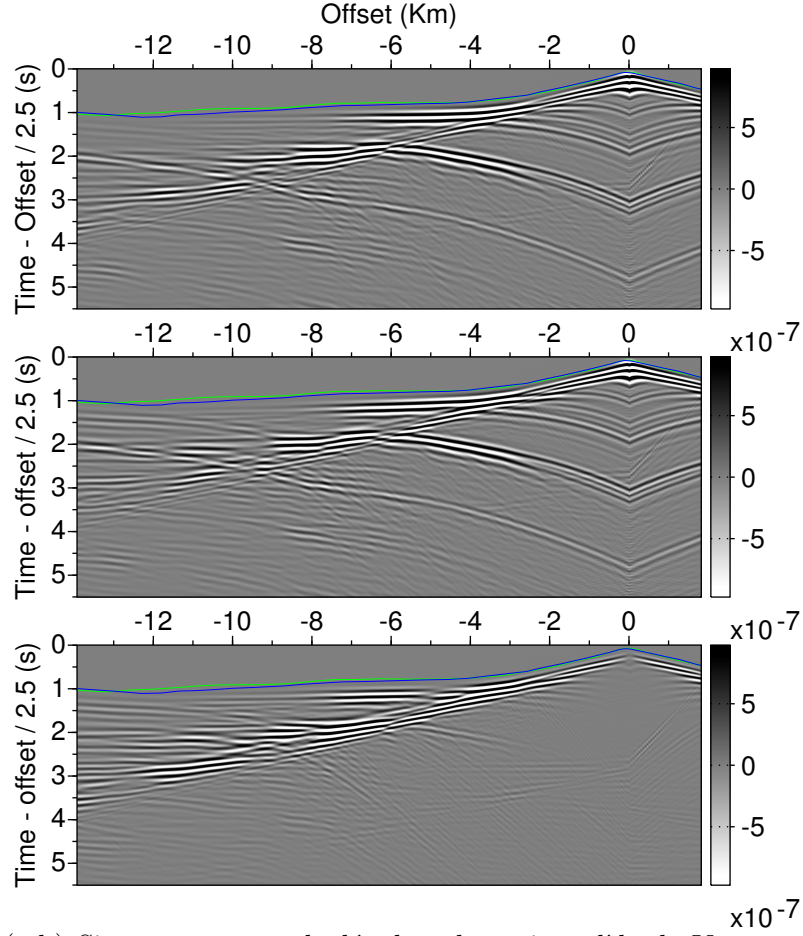


FIGURE 3.24 – (a-b) Sismogrammes calculés dans le vrai modèle de V_p , en utilisant comme paramètres de Thomsen (a) les vrais modèles et (b) les versions lissées des vrais modèles. (c) Sismogrammes résiduels associés obtenus par la différence entre (a) et (b). Dans (a-c) sont superposés les pointés manuels des temps de premières arrivées pour une modélisation où sont considérés les paramètres de Thomsen exacts (en vert) et lissés (en bleu).

considéré car elles peuvent être très sensibles au bruit ou à l'éclairage fourni par l'acquisition par exemple. Ces tests, aux configurations relativement réalistes, peuvent cependant renseigner sur la manière dont les effets se compensent ou se conjuguent. Comme il est mentionné dans l'article sur l'empreinte de l'anisotropie, le modèle horizontal de vitesse déduit des paramètres de Thomsen fait une erreur de l'ordre de 0.1 s sur l'accord avec les premières arrivées au niveau des offsets intermédiaires, soulignant que les données de réflexion ne sont pas optimales pour reconstruire les vitesses horizontales. Il était donc intéressant de savoir dans quelle mesure les imprécisions cinématiques de cette nature affectent l'inversion des formes d'ondes, basées sur l'exploitation de données petits et grands angles.

Fréquence initiale	Préconditionnement par amortissement	Pourcentage d'erreur relatif ξ (%)					
		Isotrope (1)	Approximation Anisotrope		$\xi_{2,aniso}$ — $\xi_{1,iso}$	$\xi_{3,aniso}$ — $\xi_{1,iso}$	$\xi_{3,aniso}$ — $\xi_{2,aniso}$
			(2) (ϵ, δ) vrais	(3) (ϵ, δ) lissés			
2 Hz	non ($\xi_{no-dp-2Hz}$)	4.39	4.70	5.00	0.31	0.61	0.3
	oui (ξ_{dp-2Hz})	3.84	4.24	4.84	0.4	1	0.6
	$\xi_{no-dp-2Hz} - \xi_{dp-2Hz}$	0.55	0.46	0.16			
3.5 Hz	non ($\xi_{no-dp-3.5Hz}$)	5.75	6.02	6.77	0.27	1.02	0.75
	oui ($\xi_{dp-3.5Hz}$)	4.62	4.99	5.8	0.37	1.18	0.81
	$\xi_{no-dp-3.5Hz} - \xi_{dp-3.5Hz}$	1.13	1.03	0.97			
	$\xi_{no-dp-3.5Hz} - \xi_{no-dp-2Hz}$	1.36	1.32	1.77			
	$\xi_{dp-3.5Hz} - \xi_{dp-2Hz}$	0.78	0.75	0.96			

TABLE 3.2 – Table donnant les pourcentages d'erreur associés aux modèles issus d'inversions isotrope et anisotrope, avec et sans amortissement, lorsque la fréquence initiale est de 2 Hz ou 3.5 Hz. Le cas anisotrope dit (2) est celui où les paramètres de Thomsen sont exacts, alors que dans le cas (3) ce sont des versions lissées de ces paramètres.

3.3.3 Sensibilité de l'inversion aux imprécisions de la partie superficielle du modèle initial

Alors que j'ai présenté sur les Figures 3.22 et 3.23 différents modèles obtenus en utilisant une fréquence initiale de 2 Hz, je montre maintenant sur les Figures 3.25 et 3.26 des modèles et leurs logs issus d'une inversion commençant à la fréquence de 3.5 Hz, et n'utilisant pas d'amortissement temporel. Le premier modèle de la Figure 3.25a est obtenu, tout comme pour les tests de la partie précédente, en partant d'un modèle initial calculé par lissage du vrai modèle, en utilisant une longueur de corrélation horizontale et verticale constante de 500 m. De plus, une version lissée des paramètres de Thomsen est utilisée pour la modélisation. Nous remarquons la présence à 700 m de profondeur d'un artefact basse vitesse relativement plat d'une centaine de mètres d'épaisseur. Cet artefact n'est pas visible sur le modèle obtenu avec une fréquence initiale de 2 Hz (Figure 3.22c) grâce aux bas nombres d'ondes reconstruits par les basses fréquences. Pour éliminer cet artefact, j'ai amélioré la partie superficielle du modèle initial de V_{P0} , en utilisant pour le lissage une longueur de corrélation verticale plus faible en surface qu'en profondeur, ce qui de plus simule la variation de résolution avec la profondeur des modèles obtenus par tomographie des temps de première arrivée. Pour cela, nous définissons z_1 et z_2 les profondeurs en haut et en bas du modèle pour lesquelles nous voulons que τ prenne les valeurs $\tau(z_1) = \tau_1$ et $\tau(z_2) = \tau_2$. Nous choisissons une loi puissance pour $\tau(z)$ de sorte que : $\tau(z) = \tau_0(1 + az^n)$, où la puissance n est choisie égale à 0.4. Les constantes τ_0 et a sont définies par

$$\tau_0 = \frac{\tau_1 - \tau_2 \left(\frac{z_1}{z_2}\right)^n}{1 - \left(\frac{z_1}{z_2}\right)^n} \quad \text{et} \quad a = \frac{1}{z_2^n} \left(\frac{\tau_2}{\tau_0} - 1\right). \quad (3.10)$$

Les différentes lois utilisées sont représentées sur la Figure 3.25e. Nous pouvons voir sur les logs de la Figure 3.26 comment la longueur de corrélation verticale τ_1 en surface affecte la précision du modèle initial (courbe bleue).

Tout d'abord, nous utilisons une longueur de corrélation verticale variant de $\tau_1 = 350m$ sur le fond marin à $\tau_2 = 500m$ à la base du modèle, tout en gardant une longueur de corrélation horizontale constante de 500 m. Nous observons sur la Figure 3.22b que l'artefact est toujours présent, bien que son empreinte soit plus locale (au centre du modèle). Si nous refaisons ce test en considérant dans la modélisation les vrais paramètres de Thomsen et non leur version lissée, nous voyons que les artefacts décrits dans les modèles précédents ont disparu (Figure 3.22c). Si nous continuons à améliorer la résolution de la partie superficielle du modèle initial en utilisant une valeur de $\tau_1 = 250m$, en utilisant la version lissée des paramètres de Thomsen, nous constatons que là aussi les artefacts décrits ont disparu (Figure 3.22d). La précision du modèle initial en surface est d'autant plus importante que c'est à ces profondeurs que les ondes réfractées se propagent, ondes pour lesquelles l'imprécision cinématique est la plus forte en considérant les paramètres de Thomsen lissés.

3.3.4 Sensibilité à la méthode d'optimisation et à la régularisation

Jusqu'à présent, j'ai utilisé une fonctionnelle dépendant uniquement du résidu des données, combinée à une régularisation consistant à lisser le gradient et le Hessien (régularisation 1). L'inversion des données de la ligne 29 dans cette configuration aboutit au modèle représenté

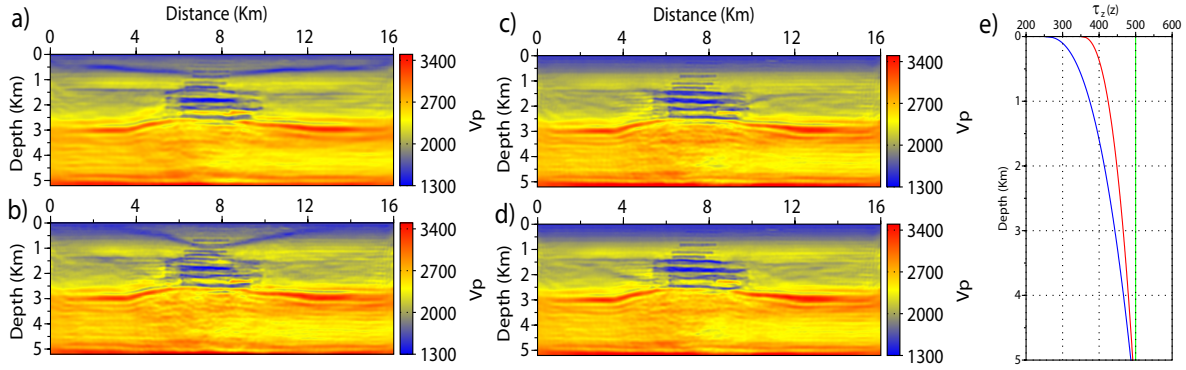


FIGURE 3.25 – Influence de la qualité du modèle initial sur les résultats. Modèles obtenus en utilisant la version lissée (a,b,d) et exacte (c) des paramètres de Thomsen, maintenus fixes au cours de l'inversion. Les modèles initiaux ont été obtenus par lissage gaussien, en utilisant une longueur de corrélation verticale constante de 500 m (a), variant de 350 m à 500 m de la surface au bas du modèle (b,c), et variant de 250 m à 500 m en (d). La fréquence initiale de ces tests est 3.5 Hz. (e) Loi des longueurs de corrélation verticales utilisées dans les différents tests en (a-d), lorsque τ_1 vaut 500 m, 350 m et 250 m (courbes verte, rouge et bleu respectivement).

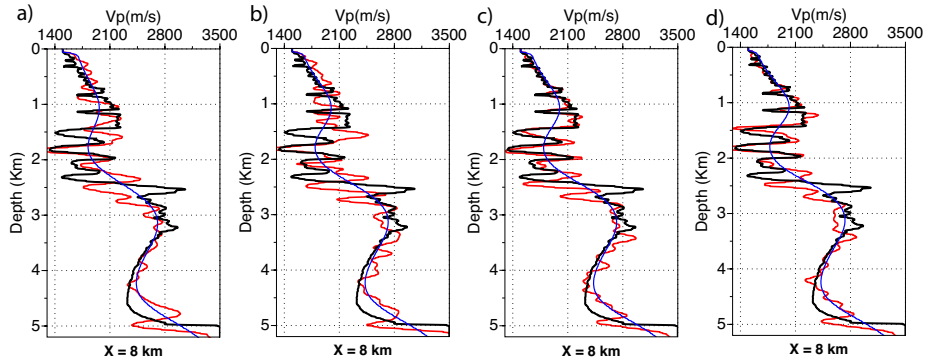


FIGURE 3.26 – (a-d) Logs extraits du centre des modèles représentés sur la Figure 3.25(a-d). Notez les différences de précision du modèle initial (courbes bleues) en subsurface.

sur la Figure 3.27a, dont nous constatons le caractère relativement lisse. Les autres caractéristiques de l'inversion sont identiques à celles utilisées pour l'application à la ligne 21 dont les résultats sont présentés dans l'article de la partie 3.2.1. Si maintenant nous utilisons l'algorithme d'optimisation L-BFGS au lieu du gradient conjugué, nous obtenons le modèle de la Figure 3.27b, qui s'avère être très bruité. Les tests sur le modèle synthétique de Valhall dans la même configuration n'ayant pas ce caractère bruité, nous en déduisons que le L-BFGS essaye de prédire le bruit des données réelles en introduisant des artefacts dans le modèle.

Pour minimiser l'empreinte de ce bruit dans les résultats, j'ai utilisé une régularisation de type Tikhonov (régularisation 2) qui avait été implémentée par d'autres membres du groupe SEISCOPE. Le terme dans l'espace des modèles rajouté à la fonction coût (Equation 1.65 du chapitre 1) est un terme de rappel pour que le modèle trouvé reste proche d'un modèle a priori. La difficulté réside dans le choix, éventuellement dynamique, du paramètre de pondération λ

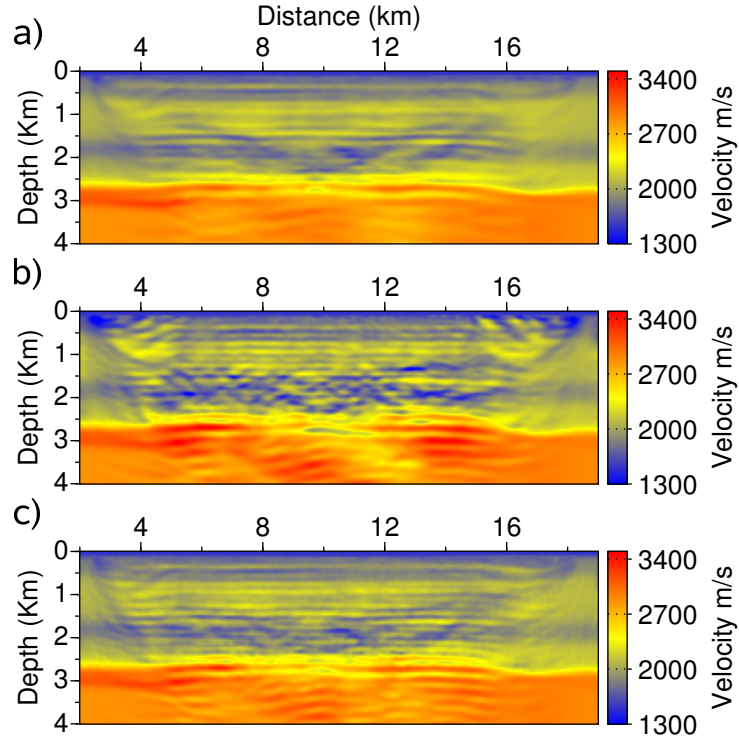


FIGURE 3.27 – Sensibilité de l'inversion des formes d'ondes à la méthode d'optimisation et à la régularisation. Modèles obtenus par inversion de la ligne 29 en utilisant une régularisation consistant à lisser le modèle de perturbation (a-b), associé à la méthode du gradient conjugué (a) et la méthode du L-BFGS (b). (c) Modèle obtenu avec le L-BFGS et une régularisation de type Tikhonov.

(Equation 1.65) qui contrôle le poids respectif des deux termes de la fonctionnelle exprimés respectivement dans l'espace des données et des modèles. Lors de l'application de cette régularisation au champ de Valhall, le modèle a priori est régénéré à chaque groupe de fréquences et est égal au modèle de départ de l'inversion du groupe de fréquences courant (c'est-à-dire le modèle final issu du groupe précédent). Comparativement à une approche utilisant un modèle a priori unique (le modèle initial du premier groupe de fréquences), l'approche dynamique que j'ai utilisée permet de reconstruire des modèles mieux résolus à hautes fréquences. Le modèle a priori est par ailleurs maintenu fixe lors des itérations de l'inversion du groupe de fréquences, ce qui garantit que la définition de la fonctionnelle ne change pas au cours des itérations, hypothèse requise par le L-BFGS qui construit récursivement le produit de l'inverse du hessien à partir des valeurs des gradients et des modèles de perturbations des itérations précédentes. La constante d'amortissement λ (Equation 1.65) a été trouvée par essai erreur de sorte que le modèle final reconstruit (Figure 3.27c) soit d'une résolution comparable à celui obtenu en utilisant le gradient conjugué et la régularisation précédente. Nous avons trouvé une valeur pour λ_{V_p} égale à 4.10^{-8} , pour laquelle il s'est avéré que le gradient de la fonctionnelle dans l'espace des données est environ (selon les itérations) 10 fois supérieur au gradient de la fonctionnelle dans l'espace du modèle. Par ailleurs, le modèle d'amortissement σ , qui est l'une des variables composant la

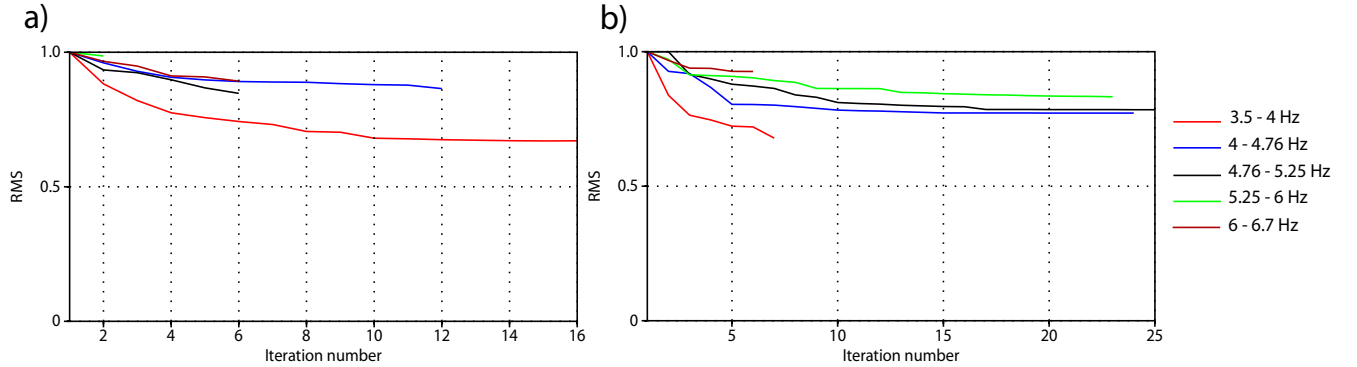


FIGURE 3.28 – Décroissances des fonctions coût en fonction du nombre d'itérations, pour les cinq groupes de fréquence (courbes colorées), pour les tests obtenus en utilisant le gradient conjugué et la régularisation 1 (a), et le L-BFGS avec la régularisation 2 (b). Notez que le nombre d'itérations maximum est différent pour les deux figures.

matrice de pondération des modèles (\mathbf{W}_m , dans l'équation 6.6), exprime l'incertitude que l'on a du paramètre. J'ai considéré qu'une valeur réaliste d'incertitude pouvait être prise égale à 10 % de la valeur moyenne du paramètre, soit pour V_p une valeur de 200 m/s. J'ai initialement essayé la régularisation 2 avec un modèle homogène de σ , mais je me suis rendu compte, en particulier pour la ligne 21, que le modèle de V_p avait tendance à reconstruire de trop hautes vitesses dans les 500 premiers mètres, peut-être en raison de la proximité avec les sources ou à cause du bruit de plateforme. Pour stabiliser l'inversion dans cette zone, j'ai utilisé un modèle d'amortissement σ variable avec la profondeur et homogène latéralement. Le paramètre σ est défini égal à 20 m/s dans les 100 premiers mètres, puis suit une loi linéaire pour atteindre 200 m/s à 400 m de profondeur, puis reste égal à 200 m/s en dessous.

Les modèles de la ligne 29 représentés sur la Figure 3.27a et c sont similaires, bien que le modèle obtenu avec le L-BFGS et la régularisation Tikhonov paraisse un peu moins lisse, et plus perturbé en profondeur que celui obtenu avec le gradient conjugué et la régularisation 1. Les mêmes conclusions peuvent être tirées pour la ligne 21 de la comparaison des modèles et des logs obtenus avec ces deux procédures (comparez les Figures 3.6a et 3.7a avec les Figures 3.29a et 3.30a). Par ailleurs, je montre sur la Figure 3.28 l'évolution des fonctions coût avec les itérations pour chacun des cinq groupes de fréquences des deux tests de la ligne 29 susmentionnés. Il apparaît que mis à part pour le dernier groupe, le test avec L-BFGS aboutit à des fonctions coût inférieures ou égales à celles obtenues avec le gradient conjugué. Cette plus forte décroissance est permise notamment par un nombre plus important d'itérations. Dans la suite de la thèse, l'ensemble des tests seront obtenus avec le L-BFGS et la régularisation Tikhonov.

3.3.5 Sensibilité aux normes sur données réelles

La norme L2 repose sur l'hypothèse d'un bruit de distribution gaussienne dans l'espace des données, et son expression implique que les résidus de fortes amplitudes prennent une part importante dans la fonction coût totale, rendant la norme plus sensible au bruit que la norme L1. La sensibilité de l'inversion des formes d'onde élastique à différentes normes a été étudiée par Brossier *et al.* (2010c). Il fut montré que la norme L1 rend l'inversion moins sensible aux bruits d'amplitude aberrante ('outlier'), car la construction du gradient se fait en

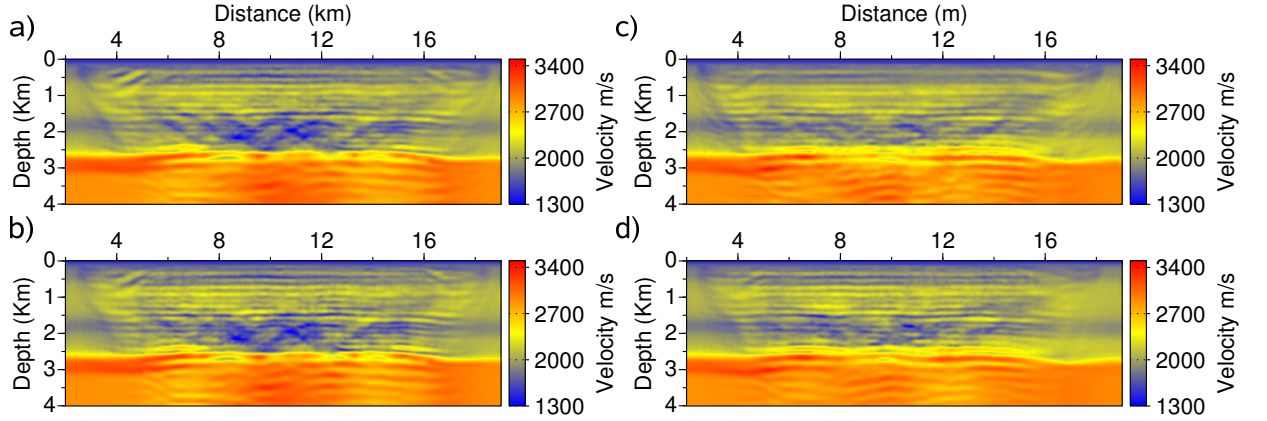


FIGURE 3.29 – Résultats d'inversion pour les lignes 21 (a-b) et 29 (c-d) avec la norme L1 (a,c) et L2 (b,d). La différence de résultats entre les deux normes est légèrement plus forte avec la ligne 29.

rétropropageant les résidus normalisés par leur module, alors que cette normalisation n'est pas appliquée avec la norme L2. La norme L2 peut cependant fournir de bons résultats en présence d'un bruit blanc en augmentant la redondance des données via un raffinement des fréquences inversées, au détriment du coût numérique. L'ensemble des tests sur données réelles se sont donc faits avec une norme L1 et en considérant des groupes de fréquences se recouvrant, mais il m'a semblé intéressant de voir l'influence de la norme sur notre application. Je montre sur les Figures 3.29 et 3.30 les résultats obtenus pour les lignes 21 et 29 avec la norme L1 et L2, en utilisant la régularisation 2 et le L-BFGS. Les modèles obtenus avec les deux normes sont en première approximation très proches, d'une part car la sensibilité de l'inversion aux normes est plus visible sur les modèles élastiques que sur les modèles acoustiques (Brossier *et al.*, 2010c), et peut-être car l'inversion est moins affectée par le bruit grâce aux groupes de fréquences utilisés. Cependant un examen détaillé révèle que les perturbations de vitesse sont un peu plus développées avec la norme L2, comme pour le réflecteur à 3.5 km de profondeur avec la ligne 21, ou au niveau du gaz et des oscillations verticales de vitesse dans les premiers 1.5 km de profondeur avec la ligne 29. Cela pourrait être lié au poids plus important de la fonctionnelle dans l'espace des données comparativement au terme régularisant lorsque la norme L2 est utilisée au lieu de la norme L1. Il faut donc adapter le terme d'amortissement λ à la norme utilisée. On note aussi une sensibilité aux normes légèrement plus grande pour la ligne 29. Il est possible qu'il subsiste dans ce jeu de données une plus grande part de données aberrantes malgré le contrôle qualité. L'autre raison pourrait être liée à une couleur du bruit différente pour les deux lignes (blanc ou autre), ce qui est plausible sachant que la ligne 21 est plus affectée que la ligne 29 par le bruit de plateforme.

3.4 Conclusion du chapitre

J'ai montré de manière empirique l'influence de l'anisotropie sur une inversion isotrope, se caractérisant par un biais dans la reconstruction des vitesses et un approfondissement des

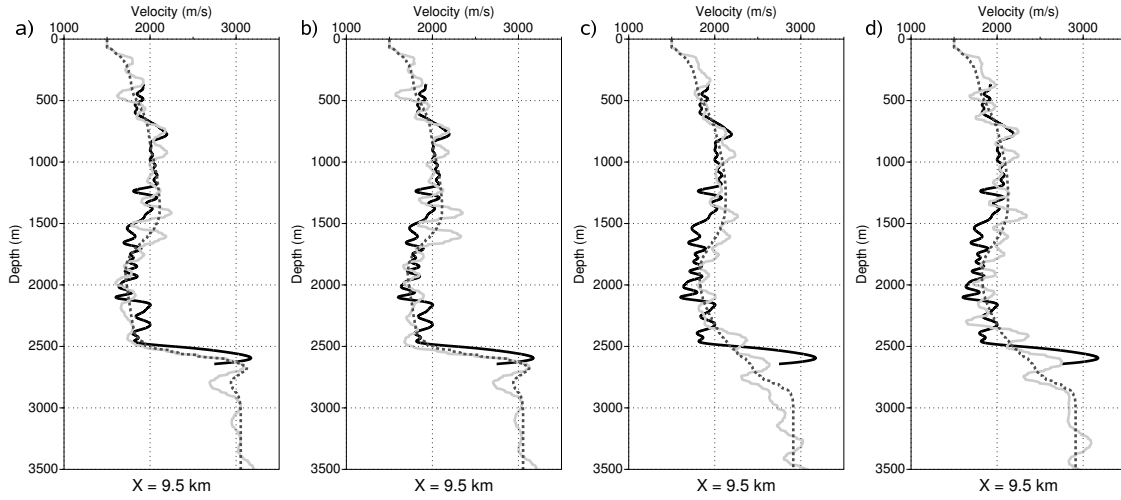


FIGURE 3.30 – Logs extraits des modèles de la Figure 3.29(a-d) (courbes grises), où nous testons la norme L1 (a,c) et L2 (b,d) pour les lignes 21 (a-b) et 29 (c-d). Les courbes noires continues et pointillées représentent le log de puit filtré localisé sur la ligne 21, et le profil extrait du modèle initial de vitesse vertical.

réflecteurs. Sur un cas synthétique isotrope j'ai illustré les non-linéarités associées aux multiples sur la surface libre et aux très longues propagations, tout en montrant que l'utilisation d'amortissements temporels pouvait réduire cette non-linéarité. J'ai montré qu'une inversion anisotrope avec des paramètres de Thomsen exacts donnait un résultat très légèrement moins bon qu'une inversion isotrope, peut-être en raison d'un problème moins bien posé (présence d'ondes S dans le champ d'ondes acoustique anisotrope) ou d'erreurs numériques plus importantes en anisotrope. Par ailleurs, j'ai quantifié l'impact des imprécisions des paramètres de Thomsen (lissés) sur l'inversion, de la fréquence initiale et du préconditionnement des données. Il s'est avéré que la stratégie des amortissements temporels est moins efficace lorsque les paramètres de Thomsen sont lissés, potentiellement à cause de la reconstruction hiérarchique de la reconstruction. Un cas synthétique nous a révélé que la précision de la partie superficielle du modèle initial de V_P s'est avérée être primordiale pour empêcher la création d'artefacts en surface liés à l'imprécision des paramètres de Thomsen.

Nous avons montré sur cas réel comment le L-BFGS associé à une régularisation fondée sur un simple lissage des modèles, échouait à fournir une image propre d'un milieu réel à cause du bruit sismique. Puis comment la régularisation Tikhonov avec une fonction coût composée de termes dans l'espace des modèles pouvait stabiliser l'inversion. Enfin, nous avons relevé peu de différences en terme de modèles obtenus avec la norme L1 et L2 pour les deux lignes de données traitées, les différences subsistant pouvant être attribuées au poids différent du terme de données par rapport à celui de modèle dans la fonction coût, et au bruit des données.

Chapitre 4

La stéréotomographie comme outil de construction d'un modèle initial pour l'inversion des formes d'ondes

Sommaire

4.0.1	Présentation de l'article	139
4.0.2	Résumé	140
4.0.3	Introduction	140
4.0.4	Méthodes	143
4.0.5	Cas synthétique de Valhall	147
4.0.6	Application au champ de Valhall	159
4.0.7	Conclusion	166
4.0.8	Appendice A : Régularisation d'une inversion aux classes multiples de données et de paramètres	172
4.0.9	Conclusion de l'article	173

4.0.1 Présentation de l'article

La précision d'un modèle de vitesse initial et sa capacité à prédire les arrivées utilisées par l'inversion des formes d'ondes est primordiale au succès de la méthode, en particulier en l'absence de basses fréquences dans les données. Bien que BP nous ait fourni un macromodèle de vitesse initial satisfaisant à notre application de Valhall, construit par tomographie en réflexion anisotrope, nous voulions mettre en avant une méthode d'imagerie plus automatique, requérant moins de *savoir-faire* dans la manière de conduire l'inversion ou dans l'interprétation des arrivées à pointer. Nous nous sommes alors intéressés à la méthode de stéréotomographie, qui tire partie du pointé semi-automatique des données dans le cube des données avant-sommation, ce qui permet de considérer une grande quantité de pointés. De plus, cette méthode conventionnellement basée sur l'exploitation des ondes réfléchies, est naturellement plus résolvante que la tomographie de temps de première arrivée. Cependant, en raison de la restriction conventionnelle de la stéréotomographie à l'utilisation des réflexions courts offsets, nous avons voulu étendre cette méthode à la prise en compte des premières arrivées.

**Building starting models for full waveform inversion from
wide-aperture data by stereotomography**

V. Prioux, G. Lambaré, S. Operto, J. Virieux

Accepted to Geophysical Prospecting in January 2012

4.0.2 Summary

Building an accurate initial velocity model for full waveform inversion (FWI) is a key issue to guarantee convergence of full waveform inversion towards the global minimum of the misfit function. In this study, we assess joint refraction and reflection stereotomography as a tool to build a reliable starting model for frequency-domain full waveform inversion from long-offset (*i.e.*, wide-aperture) data. Stereotomography is a slope tomographic method that is based on the inversion of traveltimes and slopes of locally-coherent events in the data cube. One advantage of stereotomography compared to conventional traveltime reflection tomography is the semi-automatic picking procedure of locally-coherent events, which is easier than the picking of continuous events, and can lead to a higher density of picks. While conventional applications of stereotomography only consider short-offset reflected waves, we assess the benefits provided by the joint inversion of reflected and refracted arrivals. Introduction of the refracted waves allows the construction of a starting model that kinematically fits the first arrivals, a necessary requirement for full waveform inversion. In a similar way to frequency-domain full waveform inversion, we design a multiscale approach of stereotomography, which proceeds hierarchically from the wide-aperture to the short-aperture components of the data, to reduce the nonlinearity of the stereotomographic inversion of long-offset data. This workflow that combines stereotomography and full waveform inversion is applied to synthetic and real data case studies for the Valhall oil-field target. The synthetic results show that the joint refraction and reflection stereotomography for a 24-km maximum offset data set provides a more reliable initial model for full waveform inversion than reflection stereotomography performed for a 4-km maximum offset data set, in particular in low-velocity gas layers and in the deep part of the structure below the reservoir. Application of joint stereotomography, full waveform inversion and reverse time migration to real data reveals that the FWI models and the reverse-time-migration images computed from the stereotomography model shares several features with FWI velocity models and migrated images computed from an anisotropic reflection-traveltime-tomography model, although stereotomography was performed in the isotropic approximation. Implementation of anisotropy in joint refraction and reflection stereotomography of long-offset data is a key issue to further improve the accuracy of the method.

Keywords

Stereotomography, Full Waveform Inversion, Inverse theory, seismic tomography

4.0.3 Introduction

Full waveform inversion (FWI) is an ill-posed inverse problem (*e.g.*, Tarantola, 1987; Pratt, 1999; Virieux & Operto, 2009). The ill-posedness of full waveform inversion results from noise in the data, the limited bandwidth of the seismic sources and the incomplete illumination of the subsurface provided by seismic surveys. As full waveform inversion seeks to exploit the

full wavefield, it is amenable to inversion of wide-aperture arrivals, such as diving waves and super-critical reflections recorded by long-offset acquisition geometries. These arrivals are useful for reconstruction of the large wavelengths of the subsurface and for improvement of the wave illumination in complex environments, such as subsalt and sub-basalt targets. However, such long-offset data increase the ill-posedness of full waveform inversion, because more wavelengths are propagated in the subsurface, making the optimization subject to cycle skipping artifacts (Sirgue, 2006; Pratt, 2008). Full waveform inversion conventionally relies on local optimization methods, such as steepest-descent or Gauss-Newton methods, because of the computational cost of the multi-source seismic modeling and the large number of model parameters. In this local optimization framework, the remedies to reduce the ill-posedness of full waveform inversion and to converge towards the global minimum of the misfit function are either the recording of low frequencies or the use of an accurate initial model. Therefore, although several attempts have been made recently by the oil industry to record low frequencies in land and marine environments (Plessix *et al.*, 2010a; Soubaras & Whiting, 2011), building accurate initial models for full waveform inversion remains a topical issue (Chauris *et al.*, 2008).

Possible seismic approaches for building initial models for full waveform inversion rely on traveltimes tomography, migration-based velocity analysis (MVA) and Laplace-domain inversion, apart from other potential sources of information. The traveltimes tomography methods include first-arrival traveltimes tomography (FATT) (e.g., Toomey *et al.*, 1994; Zelt & Barton, 1998), reflection traveltimes tomography (RTT) (Woodward *et al.*, 2008) or a combination of both (Zelt & Smith, 1992; Hobro *et al.*, 2003).

As well as the non-uniqueness of the solution (Aki & Richards, 2002), the drawbacks of first-arrival traveltimes tomography are three fold : first-arrival traveltimes picking is difficult in the presence of low-velocity zones, because the rays do not refract in low-velocity layers. Moreover, these low-velocity zones are challenging to reconstruct, because the rays are attracted towards high-velocity zones, hence creating a deficit of ray illumination of the low-velocity zones. Second, long-offset fixed-spread acquisition geometries are required for sufficient ray-path coverage of deep targets. Typically, maximum offsets as large as at least four times the maximum depth of investigation are necessary (e.g., Dessa *et al.*, 2004). Third, the spatial resolution of first-arrival traveltimes tomography is low, and is of the order of the radius of the first Fresnel zone, given by $\sqrt{\lambda \times o}$, where λ is the wavelength and o is the source-receiver offset (Williamson, 1991). It is still an open question whether this resolution is sufficient to provide a suitable starting model for full waveform inversion for realistic starting frequencies (typically, 4 Hz for conventional sources). First-arrival traveltimes tomography is conventionally implemented with rays, which require suitable regularization to steer the inversion towards smooth models. Alternatively, finite-frequency first-arrival traveltimes tomography relies on the sensitivity kernel computed by integration over the first-Fresnel zone and traveltimes delays computed by cross-correlation (Luo & Schuster, 1991; Marquering *et al.*, 1999; Dahlen *et al.*, 2000b). A second finite-frequency first-arrival traveltimes tomography method relies on frequency-domain waveform modeling, where complex-valued frequencies are used to compute the first-arrival traveltimes from damped monochromatic wavefields (Min & Shin, 2006; Ellefsen, 2009). First-arrival traveltimes tomography generally relies on the explicit building of the sensitivity or Fréchet derivative matrix, which can be expensive to compute and store in three-dimensional (3D) applications. Alternatively, matrix-free first-arrival traveltimes tomography based on the adjoint-state method allows very large data-sets to be handled without the explicit building of the Fréchet derivative matrix (Taillandier *et al.*, 2009).

Reflection-traveltime-tomography should provide velocity models of higher resolution than first-arrival traveltime tomography, because the Fresnel zones of the reflections are smaller than those of the first arrivals. However, reflection-traveltime-tomography is an ill-posed inverse problem because of the velocity-depth ambiguity when insufficient long-offset coverage is available (Farra & Madariaga, 1988). Moreover, the picking of traveltimes of continuous reflected events is a tedious and interpretative task, even if this task tends to be more and more automated today. As the reflection-traveltime tomography method relies on reflected events only, it does not guarantee the velocity model that results will allow first-arrival traveltimes to be matched with an error lower than half of the period of the starting frequency of full waveform inversion, in particular in anisotropic media (Prioux *et al.*, 2011) : this is a necessary requirement to prevent cycle skipping in full waveform inversion of wide-aperture data.

The MVA procedure relies on iterative applications of prestack depth-migration and velocity analysis. The governing idea of MVA is the transformation of the residual move-out information picked in common image gathers (CIGs) into velocity updates (Yilmaz & Chambers, 1984) through flatness criteria. The velocity updating can be automated without picking through minimization of a differential semblance functional in the image domain (Symes & Carazzone, 1991; Chauris & Noble, 2001; Shen & Symes, 2008). Alternatively, focusing analysis in extended common image gathers can be used for velocity updates (Yang & Sava, 2011b). The drawback of these approaches is the computational cost of the iterative application of prestack depth migration and velocity analysis. As with the reflection-traveltime-tomography method, the MVA procedure does not guarantee the kinematic accuracy of the velocity models with respect to the first-arrival traveltimes.

The third category of these methods relies on Laplace-domain and Laplace-Fourier waveform inversion (Shin & Cha, 2008; Shin & Ha, 2008; Shin & Cha, 2009). Laplace-domain and Laplace-Fourier inversions are applied to seismic data that are damped in time by multiplication with a decaying exponential function. The aim of the multiplicative decaying function in the time domain is to broaden the spectrum of the original signal towards low frequencies by convolution in the frequency domain. The Laplace-domain and Laplace-Fourier inversions are implemented with frequency-domain full-waveform-inversion methods, where complex-valued frequencies are used. The imaginary part of the frequency (i.e., the Laplace constant) controls the decay in time of the exponential function, and hence the amount of information used in time. In Laplace-domain inversion, the real part of the frequency is set to zero, unlike in Laplace-Fourier inversion, where more resolving information is inverted. These approaches have shown very convincing results in realistic synthetic examples, where reliable initial models for full waveform inversion have been built from scratch (Shin & Cha, 2009). However, the extraction of the Laplace-domain information from real noisy data is challenging, because of the rapid amplitude decay of the Laplace wavefield with offsets. Note that the use of the decaying exponential functions requires the recorded data above the first arrival to be muted, and therefore requires picking of the first arrival traveltimes as for the first-arrival traveltime tomography method.

In the present study, we assess stereotomography (Billette & Lambaré, 1998; Billette *et al.*, 2003; Lambaré *et al.*, 2004a; Alerini *et al.*, 2007; Lambaré, 2008; Nag *et al.*, 2010) as a tool to build an initial model for full waveform inversion from long-offset data. Stereotomography is a ray-based slope tomographic method, where the velocity macro-model is estimated from locally coherent events that are characterized by their slopes and traveltimes in the pre-stack data cube. One motivation behind stereotomography is that semi-automatic picking of local

events is easier than that of continuous events, and hence should provide a denser dataset than the conventional reflection-traveltime-tomography method. Locally coherent events are interpreted as single-scattered events, where the position of the diffractor points can be part of the model parametrization. So far, stereotomography has been limited to the reflected wavefields recorded by short-offset acquisitions. In the present study, we assess whether stereotomography can be applied to long-offset data by joint inversion of first arrivals, long-spread reflections and short-spread reflections. Combining refraction and reflection ray paths dramatically improves the ray-path illumination and makes the updating of the kinematic information during the initial model building as consistent as possible with that used during full waveform inversion. This requires the implication of all of the aperture or scattering angles in stereotomography.

In the first part of the present study, we briefly review the basics of stereotomography and frequency-domain full waveform inversion. Then, we show the benefits provided by the joint inversion of refraction and reflection traveltimes for building reliable initial models for full waveform inversion by stereotomography, with a realistic synthetic example inspired by the case of the Valhall oil field. In particular, we compare the results of our workflow for several datasets that are characterized by different offset ranges. We propose a new hierarchical approach of stereotomography, which allows for robust inversion of long-offset data. With this hierarchical approach, we show how long-offset data in stereotomography allow us to improve the reconstruction of the subsurface at the reservoir level and below, compared to conventional reflection stereotomography. In the last section, we show a preliminary application of our workflow to real ocean-bottom cable (OBC) data from the Valhall field. We obtain a FWI model from the stereotomographic model that shares several features with previous FWI models obtained from a reflection-traveltime tomography model provided by the BP company, although the stereotomographic model developed by isotropic stereotomography is hampered by the anisotropy of the wide-aperture data.

4.0.4 Methods

4.0.4.1 Stereotomography

Data and model spaces

In the present study, we follow the stereotomographic approach described by Billette & Lambaré (1998); Billette *et al.* (2003); Lambaré *et al.* (2004a). A brief review of the method is given below. In stereotomography, each locally coherent event picked in the pre-stack data volume can be represented by the associated source and receiver positions $\mathbf{S} = (S_x, S_z)$, $\mathbf{R} = (R_x, R_z)$, its two-way traveltime T_{SR} , and two local slopes ($p_S = \partial T_{SR} / \partial S$, $p_R = \partial T_{SR} / \partial R$) that represents the horizontal component of the slowness vectors at the source and receiver positions, respectively (Billette & Lambaré, 1998) (Figure 4.1). The local event is processed as a single diffraction from a scatterer located at an unknown position \mathbf{x} in the subsurface. The traveltime T_{SR} represents the two-way traveltime of the rays that connect the source to the diffractor point and the diffractor point to the receiver. The model parameters that are reconstructed during stereotomography consist of the velocity model parametrized by cardinal cubic B-splines coefficients C and the pairs of ray segments that connect the diffractor point to the source and receiver. These ray segments are parametrized by the diffraction angles (θ_S, θ_R) and the one-way traveltimes (T_S, T_R with $T_{SR} = T_S + T_R$). In summary, the data and model spaces involve 3 data and 4 parameter classes with different units : $\mathbf{d} = \{[\mathbf{S}_i, \mathbf{R}_i], T_{SR_i}, [p_{S_i}, p_{R_i}]\}_{i=1,N}$ and

LA STÉRÉOTOMOGRAPHIE COMME OUTIL DE CONSTRUCTION D'UN MODÈLE INITIAL POUR L'INVERSION DES FORMES D'ONDES

$\mathbf{m} = \left\{ ([\mathbf{x}_i], [T_{S_i}, T_{R_i}], [\theta_{S_i}, \theta_{R_i}])_{i=1,N}, (C_j)_{j=1,M} \right\}$, respectively, where N and M denote the number of picks and velocity nodes, respectively. During modeling, the source and receiver rays are traced upwards from the diffractor point towards the surface, using the diffraction angles θ_S and θ_R . Stereotomographic optimization aims at finding stereotomographic models, *ie.* coupled velocity values and ray segment parameters, that minimize the misfit between recorded and computed stereotomographic data.

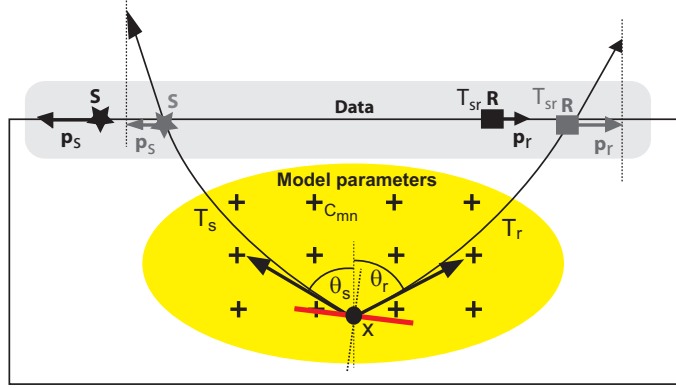


FIGURE 4.1 – Data and model parameters in stereotomography (Billette & Lambaré, 1998). The recorded data attributes inverted during stereotomography are delineated by the gray rectangle. \mathbf{S}, \mathbf{R} : source and receiver positions; T_{sr} : two-way traveltime ($\mathbf{S} \Rightarrow \mathbf{x} \Rightarrow \mathbf{R}$). $\mathbf{p}_s, \mathbf{p}_r$: horizontal components of the source and receiver slowness vectors (*i.e.*, slopes). The black and gray symbols represent measurements and modeled data. The model parameters are delineated by the yellow ellipse. C_{mn} : B-spline coefficient for P-wave velocity. The cross represents B-spline knots. T_s, T_r : one-way traveltimes, ($\mathbf{x} \Rightarrow \mathbf{S}$) and ($\mathbf{x} \Rightarrow \mathbf{R}$) ($T_{sr} = T_s + T_r$). \mathbf{x} : reflecting/diffracting point. θ_s, θ_r : source and receiver scattering angles. The red line represents a migrated facet.

The inversion algorithm

The regularized joint inversion for velocity and ray segment parameters is solved by local minimization of the least-squares misfit function \mathcal{C} using the least-squares QR (LSQR) conjugate gradient method (Paige & Saunders, 1982). As multiple classes of data and parameter are involved in the inversion, the data and model parameters must be correctly normalized to assign comparable weights to the different data and parameter classes during the inversion. A re-parametrization of the model space by weighting each parameter by a suitable norm of the corresponding column of the Fréchet derivative matrix is used to weight the contribution of each model parameter class in the inversion, following the so-called convolutional-quelling approach (Meyerholtz *et al.*, 1989). This re-parametrization is combined with a Tikhonov regularization, which is applied to the velocity parameters to guarantee the smooth reconstruction of the velocity field. The different classes of data are weighted by a prior standard deviation σ_i , where the index i denotes the class of the data, to make the data misfit vector dimensionless. The inversion scheme used in the present study is reviewed in more details in Appendix A.

Practical aspects

Stereotomography requires some preliminary steps to set the initial model parameters for the LSQR inversion. The first step provides a first estimate of the position of the diffractor points. During the first iteration of this estimation, reflecting/diffracting points are located at depth from geometrical considerations of straight rays in a homogeneous medium. In a second step, a simplified inverse problem is solved by singular value decomposition, where each pair of ray segments is individually optimized in a starting velocity model, which can be as simple as a constant gradient velocity model. Fréchet derivatives for both diffractor location and velocity updates are computed by paraxial ray tracing.

Compared to the multiscale approach presented in Billette *et al.* (2003), where the spacing between the B-spline knots is progressively refined, we introduce an additional level of hierarchy into the stereotomography that is inspired by multiscale strategies implemented in full waveform inversion of wide-aperture data : the stereotomography is subdivided into successive inversions of several sub-datasets that are characterized by an increasing short-aperture content. This allows us to proceed from the reconstruction of the long wavelengths to the short wavelengths, as shown by the relationship between the local wavenumber at the diffractor point \mathbf{x} and the aperture angle θ written as

$$\mathbf{k} = \frac{2f}{c} \cos(\theta/2) \mathbf{n}, \quad (4.1)$$

where a unit vector \mathbf{n} describes the direction of the vector obtained by summing the source and receiver slowness vectors at the position of the diffractor point, namely $\mathbf{p}_s + \mathbf{p}_r$. The frequency is denoted by f , and the local wavespeed by c (Figure 4.2). The relationship between the local wavenumber in the tomographic model and acquisition-related parameters (frequency and aperture) has been derived in the framework of various formulations of diffraction tomography (Wu & Toksöz, 1987) such as generalized inverse Radon transform (Miller *et al.*, 1987), ray+Born migration/inversion (Thierry *et al.*, 1999a) and full waveform inversion (Sirgue & Pratt, 2004). The strategy, which consists of proceeding over decreasing angles in stereotomography, is close in spirit to the multiscale approach of full waveform inversion based on hierarchical damping of seismic data (Brossier *et al.*, 2009b), which is also referred to as Laplace-Fourier inversion (Shin & Cha, 2009). As shorter apertures are involved in the inversion to improve the resolution of the imaging, the B-spline parametrization of the model is refined accordingly. This is outlined in the multiscale stereotomographic algorithm 4.1.

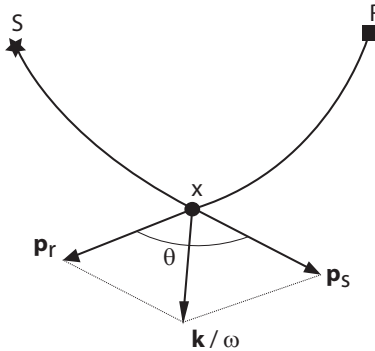


FIGURE 4.2 – Relationship between the aperture angle θ , the angular frequency ω and the wavenumber, imaged locally at the diffracting point \mathbf{x} by diffraction tomography (e.g., Thierry *et al.*, 1999b).

LA STÉRÉOTOMOGRAPHIE COMME OUTIL DE CONSTRUCTION D'UN MODÈLE INITIAL POUR L'INVERSION DES FORMES D'ONDES

It is worth noting that the early-arriving phases associated with diving waves correspond to the larger aperture angles. These diving waves generally illuminate only the upper part of the subsurface, because of incomplete offset coverage provided by conventional surface acquisition geometries. Therefore, only the shallow part of the subsurface is reconstructed during the early stages of our multiscale stereotomographic approach, which follows, therefore, a layer-stripping procedure.

Algorithm 4.1 Multiscale algorithm for stereotomography

```

1: for  $Step\_i = Step\_1$  to  $Step\_n$  do
2:   set B-spline velocity grid with suitable vertical node spacing
3:   select range of aperture angle  $\Delta\theta(i)$ 
4:   while  $iter < niter_{max}$  do
5:     Location of the events (i.e optimization of each pair of ray segments) in the current
       velocity model
6:     Computation of the Fréchet derivatives
7:     Joint inversion of rays segments and velocity model parameters, using LSQR scheme
8:   end while
9: end for

```

4.0.4.2 Frequency-domain full waveform inversion

In the present study, acoustic full waveform inversion (FWI) is performed in the frequency domain using a local optimization approach (Pratt *et al.*, 1998; Pratt, 1999).

We use FWI codes based upon frequency-domain finite-difference and discontinuous Galerkin finite-element seismic modeling for the synthetic and real data applications, respectively (Soubier *et al.*, 2009a,b; Brossier, 2011a). We use a discontinuous Galerkin finite-element seismic modeling on unstructured triangular mesh for the real data acquisition, because it provides more flexibility and accuracy for the source and receiver implementations and the description of the shallow part of the medium (Prioux *et al.*, 2011). For the synthetic case study, a finite-difference method is used because it is faster and less memory demanding (Brossier *et al.*, 2010a).

For the inversion, we use a preconditioned conjugate-gradient algorithm, where the preconditioning of the descent direction is implemented by smoothing the gradient at each iteration with a 2D Gaussian filter and dividing the gradient by the diagonal terms of the approximate Hessian. The gradient of the misfit function is computed with the adjoint-state method (Plessix, 2006; Chavent, 2009). In this framework, the i^{th} model perturbation Δm_i is given by

$$\Delta m_i = -\alpha (diag \mathbf{H}_a + \epsilon I)^{-1} \mathcal{G}_m Re \left\{ \mathbf{u}^t \left(\frac{\partial \mathbf{A}}{\partial m_i} \right)^t \mathbf{A}^{-1} \Delta \mathbf{d}^* \right\} \quad (4.2)$$

where the impedance matrix \mathbf{A} is the frequency-domain forward-modeling operator, the incident wavefield is denoted by \mathbf{u} , the radiation pattern matrix for the m_i parameter by $\partial \mathbf{A} / \partial m_i$, the Gaussian smoothing regularization operator by \mathcal{G}_m , and the data misfit vector by $\Delta \mathbf{d}$. The approximate Hessian \mathbf{H}_a is the linear part of the full Hessian, the diagonal of which corrects for the geometrical spreading of the data residuals and the partial derivative wavefields. The damping term ϵ prevents division by zeroes when the gradient is divided by the diagonal terms

of the Hessian. The step length α is computed by a parabola fitting of the misfit function along the gradient direction. The expression of the model perturbation in Eq. 4.2 is given for one source and one frequency. Contribution of multiple sources and frequencies is taken into account in Eq. 4.2 by summation of the elementary contributions of each source and frequency.

Our frequency-domain full waveform inversion follows a multiscale approach, with two levels of data preconditioning. Both of these levels aim to introduce progressively shorter wavelengths into the subsurface model. The first level of data preconditioning corresponds to successive inversions of increasing frequencies, to limit the effects of the cycle skipping ambiguity (Sirgue & Pratt, 2004). The second level of data preconditioning is implemented with decreasing exponential time damping of the data, for progressive injection of the shorter aperture components of the data and later-arriving phases (Brossier *et al.*, 2009b).

4.0.5 Synthetic Valhall case study

In the following section of the present study, we apply our workflow based upon stereotomography and full waveform inversion to the synthetic Valhall case study. The synthetic Valhall model is representative of the Valhall field located in the North Sea (Figure 4.3a). This is a shallow-water model with the sea bottom at 70 *m* in depth, which leads to strong free-surface multiples in the data. A pile of gas layers in the sedimentary cover creates a low-velocity zone between 1400 *m* and 2400 *m* in depth; a challenging reconstruction for the first-arrival traveltome tomography method. Also of note, there is a shallow channel in the sediment cover at a distance of around 11000 *m* and a depth of 300 *m*. An old graben that delineates a high-velocity interface is taken in compression at 3000 *m* in depth (Munns, 1985), with oil trapped underneath the cap rock of the anticline. As a result, the reservoir level at 2500 – 3000 *m* in depth forms a high-velocity layer below the gas layers. The deep structure is characterized by a horizontal reflector at 5000 *m* in depth that is associated with a positive velocity contrast. Between the reservoir level at 2500 – 3000 *m* in depth and the interface at 5000 *m* in depth, the velocities decrease gently with depth, without pronounced reflectors, which makes this part of the model almost seismically blind, and hence difficult to reconstruct by either first-arrival traveltome tomography or reflection-traveltime tomography.

The main refracted and reflected arrivals are labeled on a shot gather computed in the Valhall model with a two-way wave-equation finite-difference method and for a maximum offset of 14000 *m* (Figure 4.3c). The first-arrivals correspond to diving waves refracted from the sedimentary cover above the gas layers (Figure 4.3c, D1, D2, D3) and from the reservoir level (Figure 4.3c, Dr) for offsets lower and greater than 11500 *m*, respectively (Figure 4.3b). The head wave from the deep reflector at 5000 *m* in depth (Fig. 4.3c, D5) is not recorded as a first arrival for the offset range considered in this study (maximum offset 24000 *m*) because the head waves from the reservoir and the deep reflector have comparable slopes in the time-offset domain and because velocities decrease with depth between these two refractors.

4.0.5.1 Experimental set-up

Acquisition geometry and stereotomographic data set

For the stereotomography, we compute an acoustic wide-aperture data set by solving the two-way acoustic wave equation with a staggered-grid finite-difference time-domain method (Virieux, 1984). An absorbing boundary condition is set on the top of the model to avoid

LA STÉRÉOTOMOGRAPHIE COMME OUTIL DE CONSTRUCTION D'UN MODÈLE INITIAL POUR L'INVERSION DES FORMES D'ONDES

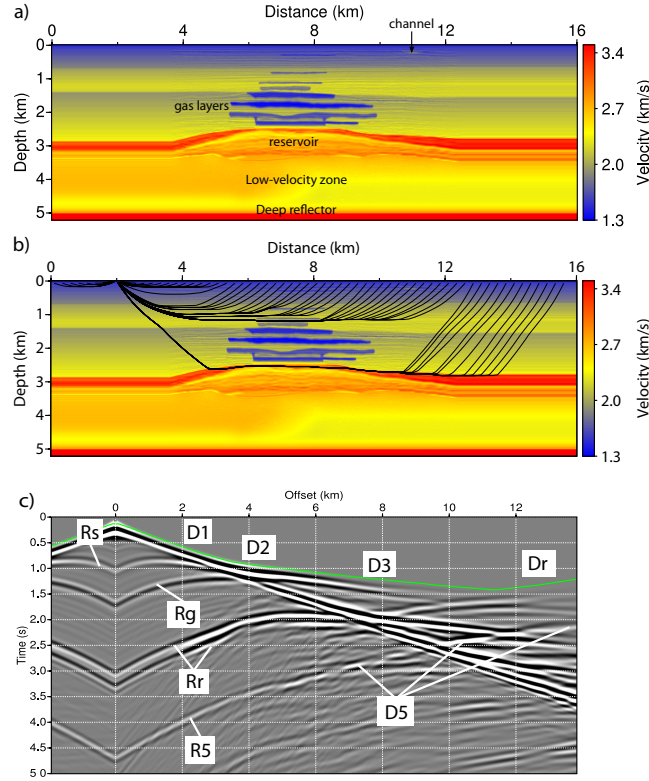


FIGURE 4.3 – Valhall synthetic model. (a) The main structural units are gas layers between 1400 *m* and 2400 *m* in depth, the reservoir at 2500 *m* in depth, a flat reflector at 5000 *m* in depth. (b) First-arrival ray tracing computed in the model shown in (a). Rays propagate in the upper structure above the gas layers and at the reservoir level. (c) Acoustic synthetic seismograms computed in the model shown in (a). The phase nomenclature is : D1, D2, D3 : diving waves in the upper structure above the gas layers. Dr : head wave from top of the reservoir. Rs : shallow reflection from the reflector located at around 700 *m* in depth. Rg : reflection from top of the gas layers. Rr : reflections from the top and bottom of the reservoir. D5, R5 : head wave and reflection from the interface located at 5000 *m* in depth.

free-surface multiples in the data, and hence to facilitate the picking during stereotomography. Therefore, we will assume in the following part of the present study that free-surface multiples can be efficiently eliminated during a pre-processing step. For the computing of a dataset with good long-offset coverage of the target (Figure 4.3), we extend the Valhall model laterally by 8000 *m* at the right and left ends of the model, and vertically by 4000 *m*, leading to a $32 \times 9 \text{ km}^2$ computational grid. The source and receiver spacings are 50 *m*, leading to 640 sources along the 32-km-long velocity model. The sources are located 6 *m* below the surface, and the receivers are on the sea bottom, hence mimicking sea-bottom acquisition. In the following, we extract several sub-datasets that are defined by their maximum offsets, to assess the footprint of the acquisition geometry on stereotomography. Although all of the B-spline knots that describe the velocity structure and the diffractor positions located in the 32-km-long velocity model are updated during stereotomography, only the 16-km-long target zone is shown (Figure 4.3).

Stereotomography data preprocessing

The data are computed with a Dirac wavelet bandpass filtered between 10 Hz and 45 Hz with a Butterworth filter. Before picking, we apply an automatic gain control (AGC) to the data, with an AGC window length of 0.5 s, for easing the picking of the late-arriving phases from the deep part of the target (Lambaré *et al.*, 2004b). We apply an external mute to remove any numerical noise before the first arrival, and an internal mute after the reflection from the deep reflector located at 5000 m in depth to remove late arrivals, which are not associated with single-scattered phases (Figure 4.4).

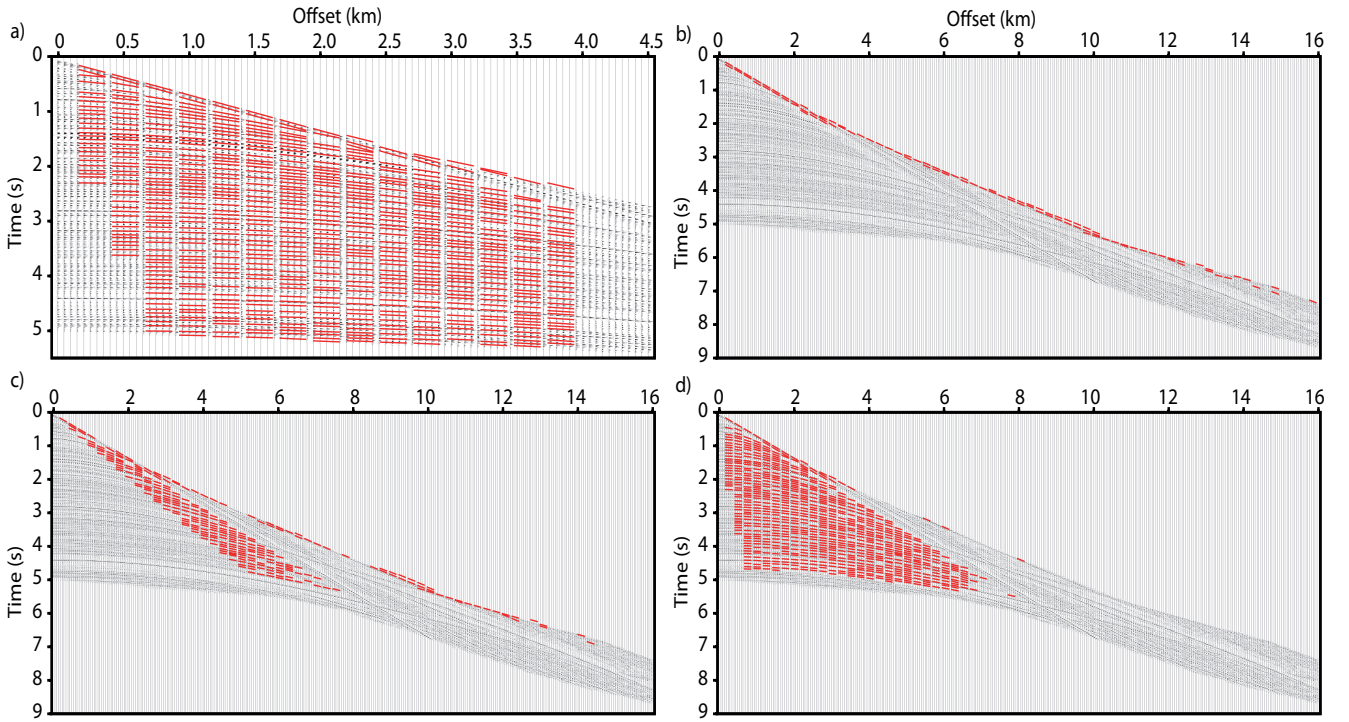


FIGURE 4.4 – Example of a shot gather with superimposed picked local slopes. The red bars represent the slopes of locally coherent events. (a) Picking is performed for a maximum offset of 4000 m. (b-d) Picking of events associated with scattering angles greater than 130° (first arrivals) (b), 80° (c) and 0° (full data set) (d).

Data picking and manual quality control

Automatic picking is performed in local slant stack panels, in both the common receiver and the common shot domains (Lambaré *et al.*, 2004b,a). Five adjacent traces on both sides of a given central trace are used to compute the semblance for a possible range of slopes. A given slope is then picked in the corresponding local slant stack panel if it satisfies different criteria of semblance. After this picking stage, a computer-assisted quality control is performed to remove picks that are identified as outliers according to several criteria, such as slope, two-way traveltimes, event coherency, and dip of migrated facets (Lambaré *et al.*, 2004b). Furthermore, the picked events that cannot be related to a diffraction point in the medium are also omitted.

Stereotomographic inversion setup

For all of the tests presented in the rest of the present study, the stereotomographic inversion is subdivided into hierarchical steps, where the vertical spacing between B-spline knots is progressively decreased from 500 *m* to 125 *m*, and the smoothing regularization is progressively relaxed accordingly. The horizontal spacing between the B-spline knots is kept at 1000 *m*, as the medium is rather tabular. Starting from a simple vertical velocity gradient model, a preliminary inversion of a few iterations is first performed to remove outliers (about 6000 outliers removed over more than 71000 picks in the first application). The quality control of the picks is repeated before each step of the multiscale stereotomography. Furthermore, locally coherent events corresponding to 'frowning' events in common image gathers computed by kinematic ray+Born migration/inversion below the gas layers were identified as internal multiples from the gas layers, and therefore removed (Lambaré *et al.*, 2004b; Lambaré & Alérini, 2005). This explains the smaller density of migrated facets in the stereotomographic models in the gas area : $X = [5000, 10500]$ *m* and $Z = [2500, 5200]$ *m* depth (Figure 4.5-a).

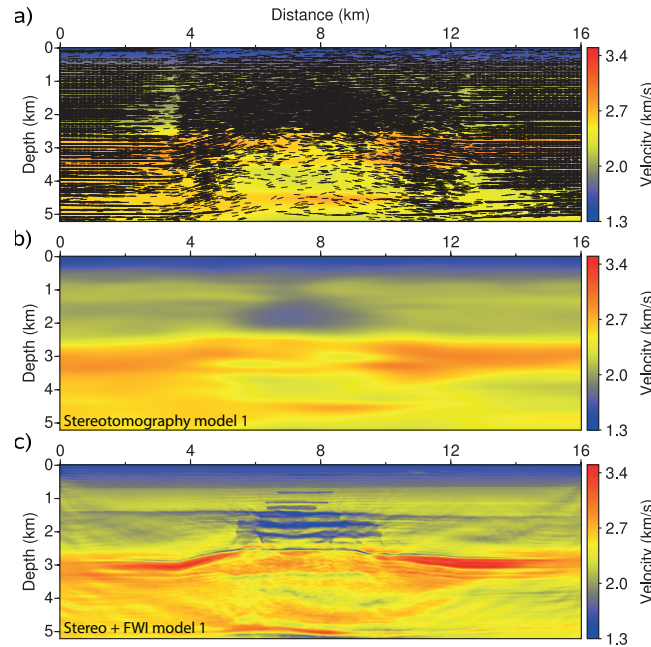


FIGURE 4.5 – Final stereotomographic model inferred from the 4-km maximum offset data set (stereotomography model 1), plotted with (a) and without (b) superimposed migrated facets. (c) FWI model obtained using the stereotomographic model shown in (a, b) as a starting model.

Full waveform inversion : acquisition geometry and inversion setup

For full waveform inversion, we consider a sea-bottom fixed-spread acquisition geometry limited to the 16-km-long target. The source and receiver spacings are 100 *m* and 50 *m*, respectively. The number of sources and receivers along the 16-km-long profile are 161 and 321, respectively. Seven monochromatic data-sets that are associated with the frequencies 4, 5, 6, 7.5, 9, 12 and 15 *Hz* are inverted sequentially. For each mono-frequency inversion, four time-damping factors of 0.5, 1, 3 and 6 *s* are successively applied to the data during the inversion. Twenty iterations are performed per frequency and time damping. For full waveform inversion, we computed the data with a free-surface boundary condition on top of the model.

Therefore, free-surface multiples are included in the inversion. The recorded and modeled data are computed with the same modeling engine that is based upon a finite-difference frequency-domain discretization of the scalar wave equation parametrized by the compressional wavespeed only. Therefore, the density is not considered during this case study. An inverse crime strategy is used, because we want to focus our analysis on the footprint of the starting model on full waveform inversion. The velocity model of Figure 4.3 was extended down to 6000 m in depth for full waveform inversion to be able to compute the reflection and the head wave from the interface located at 5000 m in depth.

4.0.5.2 Short-offset inversion

In a first application, the stereotomographic inversion is limited to reflected arrivals with a maximum offset of ± 4000 m . The 640 sources along the 32-km-long extended model are used for stereotomography. The offset coverage is laterally uniform in the target zone of Figure 4.3. The picked events (traveltimes and slopes) are superimposed on a shot gather in Figure 4.4a. The final stereotomographic velocity model (referred to as stereotomography model 1) is shown in Figure 4.5(a-b) with and without superimposed migrated facets (Billette & Lambaré, 1998; Billette *et al.*, 2003). The long-wavelength signature of the low-velocity gas layers is reconstructed, which is not the case for the low wave-numbers of the deep interface at 5000 m in depth due to insufficient offset coverage (Figure 4.6a). The final full waveform inversion (FWI) model obtained starting from the stereotomographic model is shown in Figure 4.5c. The upper structure down to the top of the gas layers at 1500 m in depth is relatively well reconstructed, including the shallow channel at a 11000 m of distance. A closer inspection of the FWI results in the middle of the model at a horizontal distance of 8000 m shows, however, an inaccurate reconstruction of the top of the reservoir at 2300 – 2700 m in depth (Figure 4.6e). This suggests that the resolution of the stereotomography model is not enough to guarantee accurate FWI reconstruction of the reservoir when a starting frequency of 4 Hz is used. The positive velocity contrast at 5000 m in depth is well delineated, but the low wave numbers of this velocity contrast are lacking. This deficit of low wavenumbers in the deep part of the FWI model results from a lack of low wavenumbers in the stereotomographic model at these depths (Figure 4.6a) and from an insufficient long-offset coverage in the dataset processed during the full waveform inversion. For completeness, a stereotomography log and the corresponding FWI velocity log near the end of the model at a horizontal distance of 14 km show some slight mispositioning in depth of the reflectors down to 3 km in depth (Figure 4.7a,c). Moreover, the signature of the reflector at 5 km in depth is fully lacking in both logs due to insufficient long-offset coverage.

4.0.5.3 Sixteen-kilometer-long offset inversion

We consider now a dataset with a maximum offset of ± 16000 m for the 640 sources deployed along the 32-km-long extended model. In this case too, the offset coverage is laterally uniform in the target zone of Figure 4.3. The picking of both the diving waves and the reflected phases is performed for stereotomography within the ± 16000 m offset range. However, we manually omitted picks in the shot gathers located in time between the first arrival and the water-layer direct arrival (Figure 4.4(c-d)). These picks are associated with multi-refracted arrivals in the shallow sedimentary layers, which cannot be taken into account by stereotomography, as this relies on the single scattering approximation. Inversion of these phases leads to unstable stereotomographic reconstructions. The corresponding picks are removed using mute laws defined

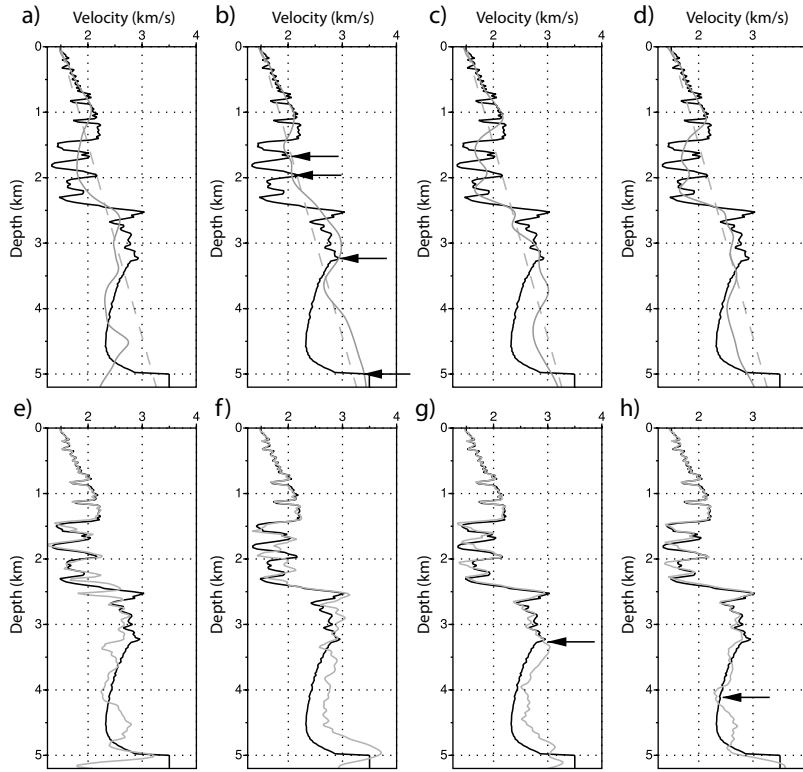


FIGURE 4.6 – Comparison between true (black line), stereotomography (a-d) and FWI (e-h) velocity profiles at 8000 *m* of distance. In (a-d), the dashed gray and the solid gray curves are taken from the initial (velocity gradient) and final stereotomography model, respectively. Maximum offset in the data during stereotomography is 4000 *m* (a), 16000 *m* (b, c), 24000 *m* (d). In (b), all the picks are inverted simultaneously by stereotomography, while a hierarchical procedure with respect to aperture angle was used in (c). (e-h) FWI models inferred from the stereotomographic models shown in (a-d). In (b), arrows show high velocity contrasts, which have been reconstructed by stereotomography. In (g, h), arrows show the depth below which velocities are overestimated due to insufficient long-offset coverage.

in the time-offset domain. The picks are superimposed on a shot gather in Figure 4.4(b, d) for the first and the reflected arrivals, respectively. Of note, the maximum offset for the reflected picks is 8000 *m*, because beyond this offset, interference between the reflected wavefields and the multi-refracted phases prevents the picking of the reflected phases.

Simultaneous inversion

In a first test, all of the picked events (the union of the picks of Figures 4.4(b, d)) are simultaneously inverted during the stereotomography. The final stereotomographic model (referred to as stereotomography model 2) and the corresponding full waveform inversion (FWI) model are shown in Figure 4.8(a-c). The FWI model of Figure 4.8c is clearly noisier than the FWI model inferred from stereotomographic model 1 (Figure 4.5c). Although the high-velocity contrast on top of the reservoir is well positioned in depth in the log of the FWI model (Figure 4.6f), the reconstruction of the low-velocities in the gas layers both in the stereotomography model and

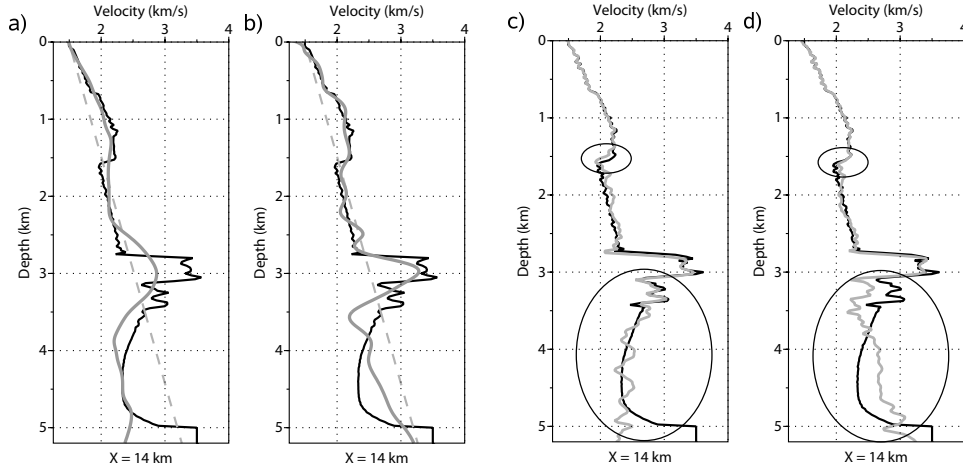


FIGURE 4.7 – Comparison between true (black line), stereotomography (a-b) and FWI (c-d) velocity profiles at 14000 *m* of distance. In (a-b), the dashed gray and the solid gray curves are taken from the initial (velocity gradient) and final stereotomography model, respectively. Maximum offset in the data during stereotomography is 4000 *m* (a) and 24000 *m* (b). (c-d) FWI models inferred from the stereotomographic models shown in (a-b). The shallow ellipse highlights the improvement of the reflector on top of the gas layers provided by the long-offset data (d). The deep ellipse highlights some artifacts in the long-wavelength reconstruction on the deep target when long-offset data are used (d). These artifacts are not shown in (c), where the deep reflector at 5 km in depth is not reconstructed at all from the short-offset data.

the FWI model is not accurate (Figure 4.6(b,f)). It is worth noting that the stereotomography appears to have been steered towards the reconstruction of the high-velocity contrasts at the expense of the low-velocity contrasts (see the arrows in the log of Figure 4.6b at 1700, 2000, 3200 and 5000 *m* in depth). These high-velocity contrasts lead to head waves in the data, which are recorded as early-arriving phases. In addition to the reflection picks, these refracted picks might have given more weight to the velocity nodes during stereotomography, which sample high velocities at the expense of those that sample low velocities. The stereotomographic and FWI inversions clearly make some attempts to detect the velocity contrast at 5000 *m* in depth. However, the velocities are overestimated in the overburden between 3500 *m* and 5000 *m* in depth (Figure 4.6(b,f)). These overestimated velocities are the footprint of the initial velocity-gradient model (Figure 4.6(a-d)), which has been poorly updated in this zone due to insufficient ray coverage. It is worth noting that these overestimated velocities are not present in the stereotomography model inferred from the reflection picks only (Figure 4.6(a,e)) : these high-velocity artifacts result from our attempt to invert first-arrival traveltimes during stereotomography, as we shall see in the following part of the present study.

Hierarchical inversion

The various inaccuracies in the model reconstruction mentioned in the previous section prompt us to implement a level of hierarchy based upon aperture selection. Five ranges of aperture angles are considered successively : $\theta > 130^\circ$ (early-arriving phases), 80° , 60° , 30° and 0° (whole data set). Picks associated with some of those classes of aperture angles are superimposed on a shot gather in Figure 4.4(b-d).

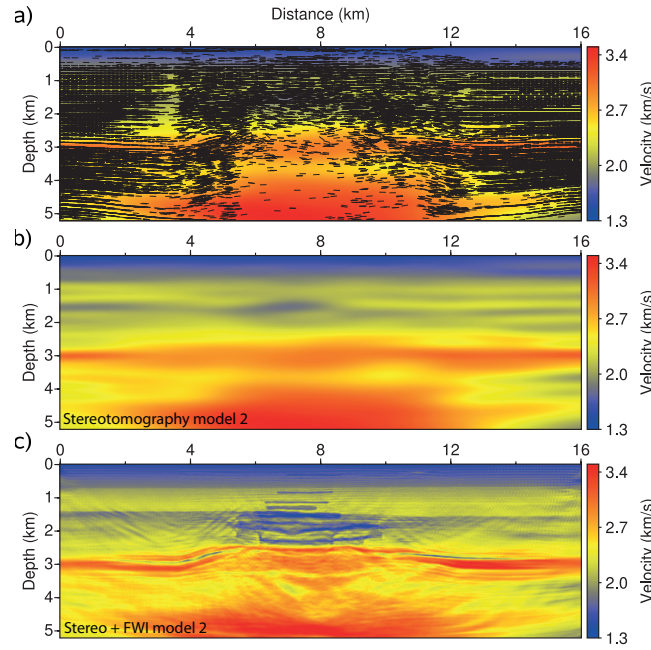


FIGURE 4.8 – Stereotomographic model obtained by simultaneous inversion of a 16-km maximum offset data set (stereotomography model 2), plotted with (a) and without (b) superimposed migrated facets. (c) FWI model inferred from the initial model shown in (a, b).

The final stereotomographic model (referred to as stereotomography model 3) and FWI velocity model inferred from this multiscale approach are shown in Figure 4.9. The corresponding vertical profiles at a distance of 8000 *m* are shown in Figure 4.6(c, g). A significant improvement in the stereotomographic model can be noted compared to the two previous tests, in particular in the gas layers where the vertical resolution is significantly improved. The FWI model shows an accurate reconstruction of the velocity structure down to the bottom of the reservoir at 3300 *m* in depth (Figure 4.6g). However, the deep part of the model remains poorly reconstructed, with overestimated velocities between 3400 *m* and 5000 *m* in depth because of insufficient long-offset coverage (Figure 4.6g, arrow). These overestimated velocities lead to an artificial deepening of the reflector located at 5000 *m* in depth (Figure 4.9c).

Figure 4.10a shows a shot gather with superimposed local slopes picked in the inverted data and computed in the final stereotomographic model 3. For the sake of clarity, these two series of slopes are plotted at the same traveltimes as those of the picked data. In Figure 4.10b, the residual slopes between the picked and modeled slopes are superimposed on the shot gather after multiplication by a factor of 25. This allows it to be shown how the slope misfit tends to increase with offset and time as more wavelengths are propagated.

4.0.5.4 Twenty-four-kilometer-long offset inversion

Further improvement to the reconstruction of the deep structure of the Valhall model needs the offset range to be increased in the acquisition up to a maximum offset of 24000 *m*. This amounts to the feeding of the inversion with additional first-arrival picks. It should be remembered here that the amount of reflection picks does not change from the 16-km to the 24-km

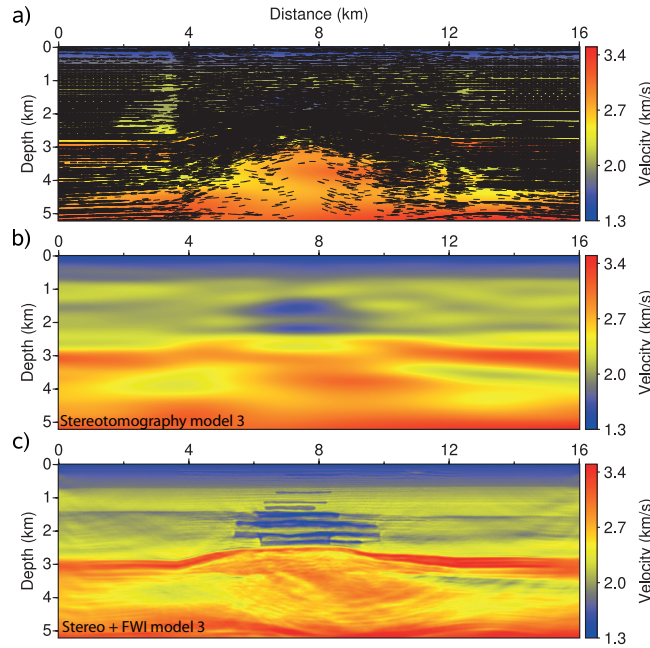


FIGURE 4.9 – Stereotomographic model obtained by hierarchical inversion of a 16-km maximum offset data set (stereotomography model 3), plotted with (a) and without (b) superimposed migrated facets. (c) FWI model inferred from the initial model shown in (a-b). Note the improvement in the stereotomographic and FWI models provided by the hierarchical approach, compared to that of Figure 4.8.

offset acquisition, because only a maximum offset of 8 km can be used to pick reflections (Figure 4.4). The velocity models at different stages of the multiscale stereotomography are shown in Figure 4.11(a-c). Note how the low-velocity zone associated with the gas layers, and the high-velocity layer on top of the reservoir at 2500 *m* in depth, progressively build up and become better resolved as we introduce picks associated with decreasing apertures angles. The FWI model inferred from the final stereotomographic model (referred to as stereotomography model 4) is shown in Figure 4.11d. The deep-velocity structure between 3200 *m* and 5000 *m* in depth has been significantly improved, due to the information provided by the long-offset first-arrival picks (Figure 4.6h, arrow). The velocity contrast at 5000 *m* in depth is well reconstructed in the middle of the profile, although the velocities between 4000 *m* and 5000 *m* in depth remain overestimated (Figure 4.6h, arrow). This again causes the mispositioning of the reflector at 5000 *m* in depth off the middle of the model in Figure 4.11d, although this mispositioning is significantly reduced compared to that shown in Figure 4.9c. It is worth noting also that the reconstruction in the gas layers has been slightly improved compared to that inferred from the 16-km-long offset data-set (compare Figures 4.6(g, h)). The stereotomography and FWI logs at 14 km of distance are shown in Figure 4.7(b,d). Comparison with the logs extracted from the stereotomography+FWI model 1 (Figure 4.7a,c) shows an improvement of the positioning in depth of the reflectors down to 3 km in depth when long offsets are involved in the inversion. It is worth noting that the attempt to reconstruct the deep reflector at 5 km in depth from long-offset rays, which sample the target from one side only, probably leads to some artifacts in the long-wavelength velocity structure between 3 km and 5 km in depth (Figure 4.7d). These artifacts are not shown in the middle of the model, where long-offset rays sample the target

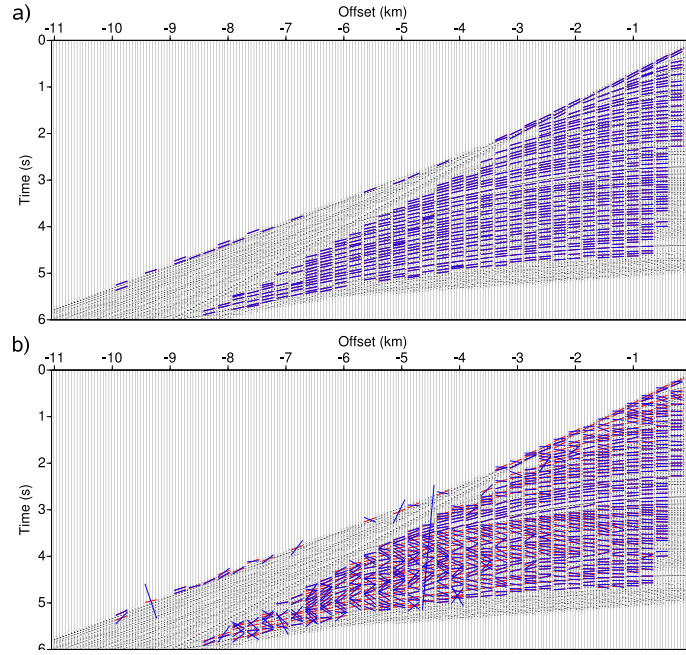


FIGURE 4.10 – (a) Shot gather with superimposed picked (red) and computed (blue) slopes. The blue slopes are computed in the stereotomographic model of Figure 4.9. The two series of slopes are plotted at the same traveltimes to facilitate the comparison. (b) The slope residuals magnified by a factor 25 are plotted in blue.

from the two sides (Figure 4.6h).

It can be asked also why the first-arrival picks have contributed to improve the imaging of the deep-velocity structure down to 5000 *m* in depth, as the deep-velocity structure between 3200 *m* and 5000 *m* in depth is a low-velocity zone where turning rays do not propagate. This can be explained by examining the rays computed for aperture angles between 175° and 185° in the stereotomographic model inferred at the end of the first stage of the 24-km maximum offset stereotomographic inversion (i.e., $\theta > 130^\circ$) (Figure 4.12). Although most of the diving rays turn in the upper structure above the gas layers, long-offset rays turn below the reservoir level down to maximum depth of around 4000 *m*. This is precisely the depth below which overestimated velocities associated with a deficit of ray coverage are shown in Figure 4.6h. Long-offset rays turn within two high-velocity gradient zones near the lateral ends of the model, which are relict from the initial-velocity gradient model. These rays allow for the sampling of the low-velocity zone located in the central part of the model by undershooting (Figure 4.12a). As short aperture angles are injected into the inversion, these positive-velocity gradient zones near the ends of the model are progressively removed as the information provided by the reflection picks contributes to focus the high-velocity contrast at the reservoir level (compare the lateral ends of the velocity models in Figure 4.11(a-c)). This explains why the footprint of the initial velocity-gradient model is not visible in the stereotomography model when only the reflection picks are used during the inversion (Figure 4.6a). The progressive removal of these velocity-gradient zones might explain why the number of first-arrival picks decreases during the multiscale stereotomography, as shorter aperture angles are injected into the inversion. Another reason might be instabilities of the ray tracing for grazing rays ($\theta = 180^\circ$) as the heterogeneity

of the stereotomographic model increases. Comparison between the results obtained with the 16-km-long and 24-km-long offset acquisitions agree that continued increases in the maximum offset range should allow us to keep on improving the reconstruction of the deepest part of the target.

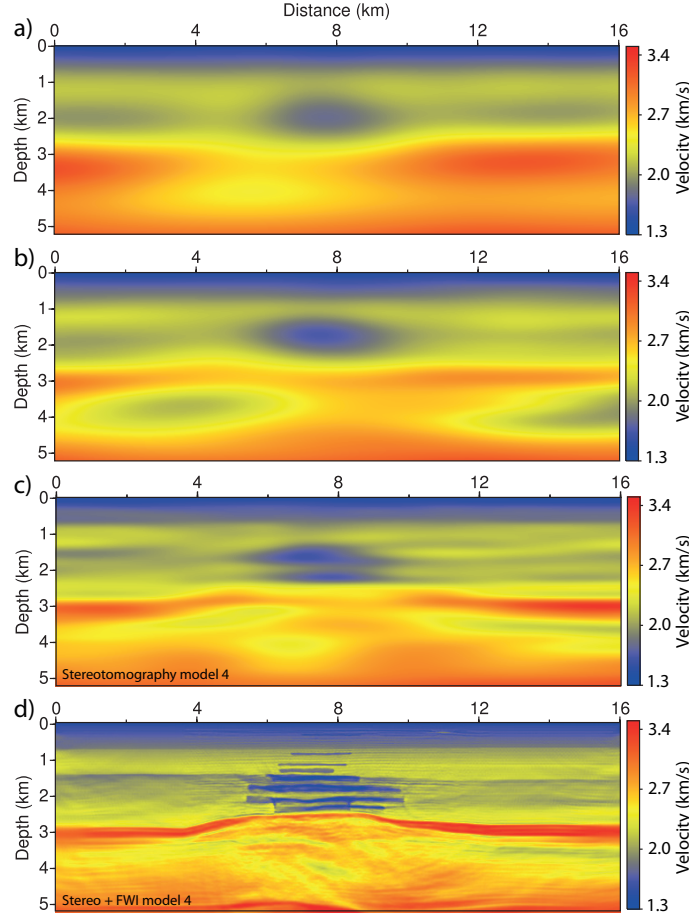


FIGURE 4.11 – Stereotomographic models obtained at different stages of the hierarchical inversion for the 24-km maximum offset data set. (a) 80° (step 2), (b) 30° (step 4), (c) 0° (step 6, all the data included) (stereotomography model 4). (d) FWI model based upon the initial model shown in (c).

To show the reliability of the ray tracing performed by stereotomography at grazing angles, we compare the ray trajectories computed by stereotomography for aperture angles between 175° and 185° and those computed with the eikonal solver of Podvin & Lecomte (1991) (Figure 4.12(b-c)). The two sets of rays show relatively good agreement. It is worth noting that the stereotomographic inversion has positioned the diffractor points at the turning point for most of the rays in Figure 4.12b, although there is an un-determination of the position of the diffractor point along the ray when the aperture angle is 180°. This probably results from the footprint of the initial stereotomographic model, which positions the initial diffractor point midway between the source and the receiver in the initial laterally-homogeneous velocity-gradient model, and hence at the turning-point position. The velocity model obtained during the first step of the multiscale stereotomography ($\theta > 130^\circ$) for the 24-km maximum-offset data-set (Figure

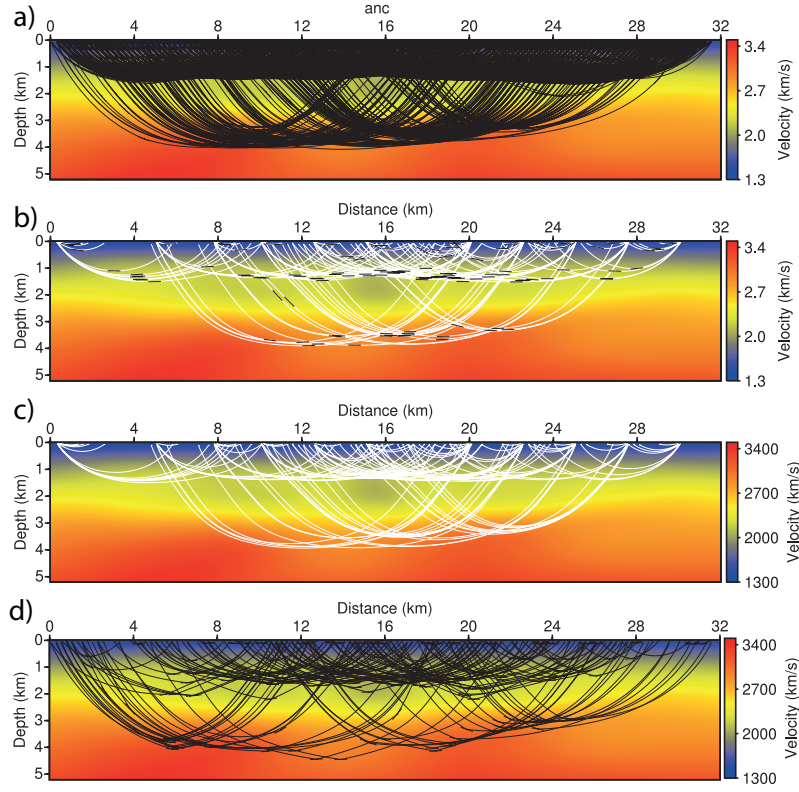


FIGURE 4.12 – (a-c) Velocity model after the first step of the hierarchical stereotomography ($\theta > 130^\circ$) for the 24-km maximum offset data set. (a) All of the rays for the aperture angles between 175° and 185° are superimposed. (b) Subset of rays shown in (a), corresponding to 11 equally-spaced shots, with their corresponding migrated facets in black. Note that most of the migrated facets are located at the turning point. (c) First-arrival rays computed with the eikonal solver of Podvin & Lecomte (1991) for the 11 shots considered in (b). Note the good match between the ray trajectories shown in (b) and (c). (d) Rays associated with aperture angles between 160° and 170° .

4.13(a-b)) is compared with the velocity model inferred from conventional first-arrival travel-time tomography (FATT) for a fixed-spread acquisition with a maximum offset of 32000 m (Figure 4.13c). The spatial resolution of these two velocity models are quite close, hence validating the relevance of stereotomography for first-arrival traveltimes inversion. The FWI model obtained using the FATT model as the initial model is shown in Figure 4.13e, 4.13f. Comparisons between the FWI models inferred from the first-arrival traveltime tomography model and the final stereotomography model 4 highlight the benefit of joint refraction/reflection traveltime tomography to build initial models for full waveform inversion (FWI) (Figures 4.6h, 4.11d, 4.13e, 4.13f).

To further assess the stereotomography models and the corresponding FWI models, we compare time-domain synthetic seismograms computed in stereotomography and FWI models 1 and 4 (Figure 4.14). The FWI model inferred from stereotomographic model 1 allows us to match the arrivals refracted and reflected from above the gas layers relatively well. However, the FWI model inferred from stereotomographic model 4 leads to a much better agreement of the arrivals refracted and reflected from below the gas layers (Figure 4.14(c, f)).

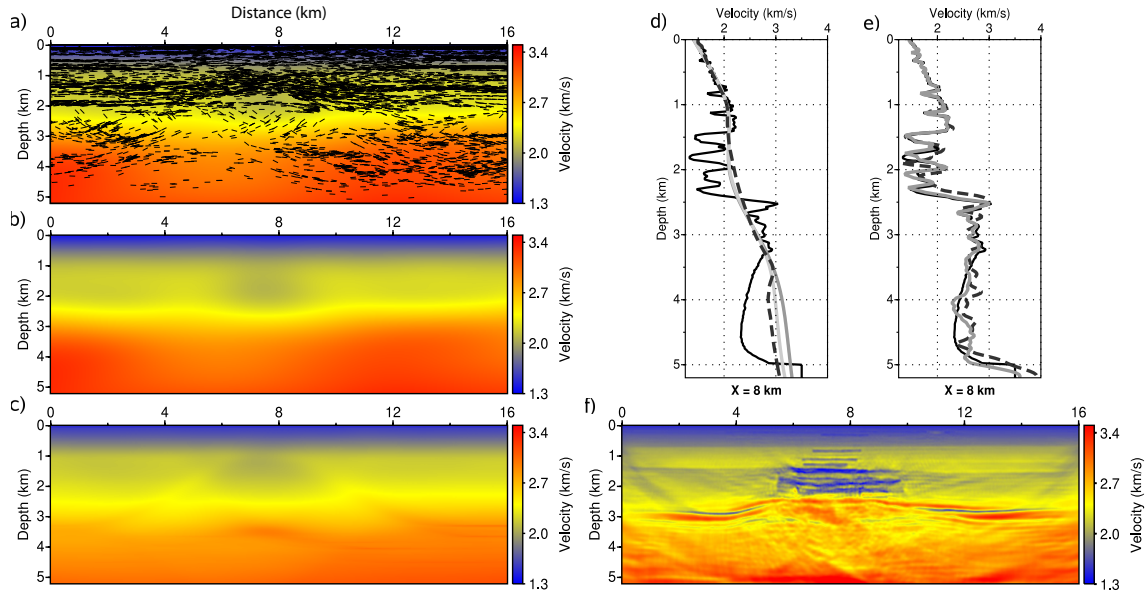


FIGURE 4.13 – (a-b) Velocity models obtained at the first stage of the stereotomography ($\theta > 130^\circ$) using a maximum offset of 24 km with (a) and without (b) superimposed migrated facets. (c) Velocity model obtained by conventional first-arrival traveltime tomography (FATT). (d) Comparisons between vertical logs at 8000 m in distance of the true velocity model (solid black line), the FATT model (black dash line), the stereotomographic model shown in (b) (light gray line) and the stereotomographic model obtained at stage 1 for the 16-km maximum offset data set (dark gray line). (e) Comparison between the vertical logs at 8000 m distance, extracted from the true velocity model (solid black line) and the FWI models inferred from the FATT model (black dashed line) and the final stereotomographic model (gray line) obtained for the 24-km maximum offset data-set. (f) FWI model obtained using the FATT model as the initial model.

4.0.6 Real Valhall case study

We now apply our workflow to an ocean-bottom-cable (OBC) data set recorded in the Valhall oil field (North Sea). In the present study, we consider only cable 21 of the 3D OBC data set (Figure 4.15). This line contains 320 shots recorded by 220 4-component receivers for a maximum offset of 13200 m. Applications of 2D acoustic anisotropic/isotropic full waveform inversion (FWI) to this OBC data set are described in Prioux *et al.* (2011), to where the reader is referred for more details. A brief overview is given here since the anisotropic FWI model developed in Prioux *et al.* (2011) will be used as a reference to appraise the anisotropic FWI model inferred from the stereotomographic model developed in the present study. Here, we will follow the same FWI experimental set-up to that in Prioux *et al.* (2011), with only the initial model used for the full waveform inversion changes. For completeness, an application of 3D acoustic isotropic full waveform inversion to the full OBC data-set was also presented by Sirgue *et al.* (2010), with high-resolution pictures of a complex network of channels in the near surface, a gas cloud (Figure 4.15), and several gas-filled fractures at depth. A starting vertical transversely isotropic (VTI) model for acoustic anisotropic full waveform inversion (referred to as the RTT model in the following) was provided by the BP company and was developed by anisotropic VTI reflection-traveltime tomography (Figures 4.16(a-b) and 4.17a,e,f). The

LA STÉRÉOTOMOGRAPHIE COMME OUTIL DE CONSTRUCTION D'UN MODÈLE INITIAL POUR L'INVERSION DES FORMES D'ONDES

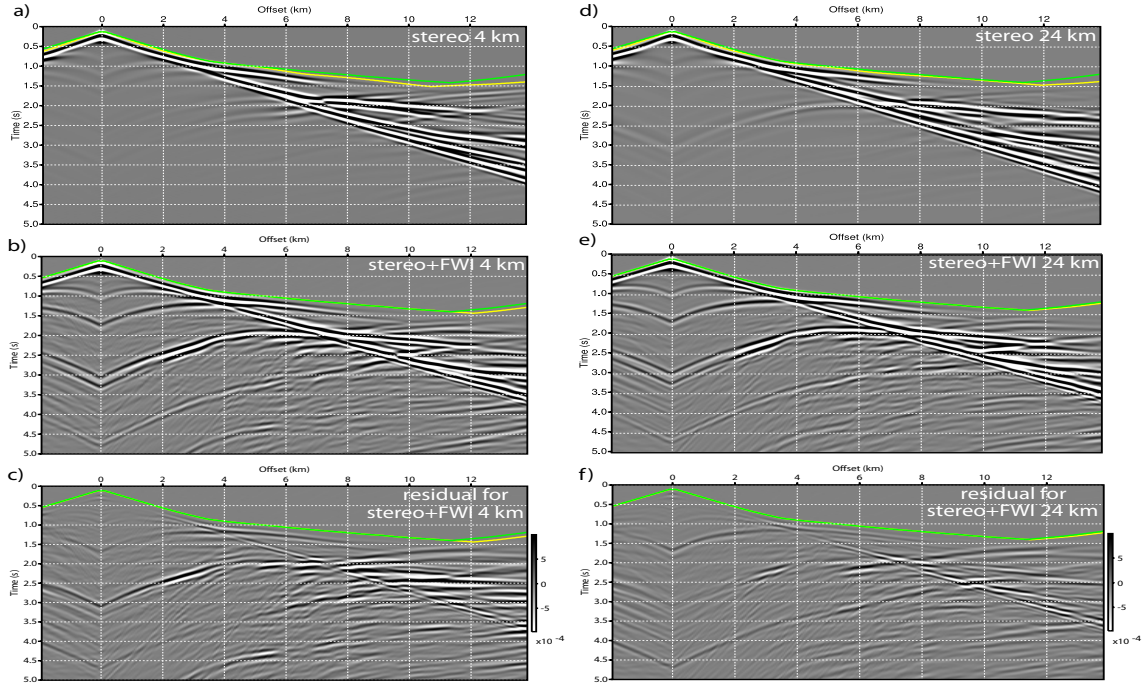


FIGURE 4.14 – Synthetic seismograms in the Valhall model. (a-b) Synthetic seismograms computed in the stereotomography models inferred from the 4-km maximum offset data set (a) and from the 24-km maximum offset dataset (d). (b-e) Synthetic seismograms computed in the FWI models obtained starting from the stereotomography model obtained from the 4-km (b) and 24-km (e) maximum offset data sets. The first-arrival traveltimes (green curve) computed in a smoothed version of the true Valhall model are superimposed along with the first-arrival traveltimes computed in the velocity models within which the synthetic seismograms were computed (yellow curve). (c) Residuals between seismograms computed in the true model (Figure 4.3c) and seismograms shown in (b). (f) Same as (c) for seismograms shown in (e).

anisotropic velocity models (either the normal-move out (NMO), the vertical or the horizontal wave speeds referred to as V_{NMO} , V_{P0} , V_h , respectively) show a low-velocity zone associated with gas layers between 1500 m and 2500 m in depth, above the reservoir level, which marks a high-velocity contrast, in a comparable way to that of the synthetic Valhall model introduced in the previous sections (Figures 4.16a). Anisotropy is significant in the Valhall data set, and it reaches a maximum of 16 % as shown by the model of the η anisotropic parameter (Alkhalifah & Tsvankin, 1995) (Figure 4.16b). The kinematic footprint of the anisotropy is illustrated by isotropic ray tracing performed in the normal move-out velocity model and the horizontal-velocity models, respectively (Figure 4.16(c-f)) : the normal-moveout velocity model allows us to match accurately the observed traveltimes of the reflections from the top and the bottom of the gas layers. In contrast, the computed first-arrival traveltimes show a significant delay at long offsets relative to the observed ones. Only the horizontal-velocity model allows us to match accurately first-arrival traveltimes, which are associated with waves traveling sub-horizontally in the subsurface. This implies that an isotropic first-arrival traveltime tomography is expected to reconstruct the horizontal velocities in the upper part of the subsurface. During anisotropic full waveform inversion performed in Prioux *et al.* (2011), the subsurface model is parametrized by the vertical velocity and the Thomsen parameters δ and ϵ (Thomsen, 1986). Only the vertical

velocity was updated (Figure 4.17b), while the δ and ϵ background models were kept fixed during full waveform inversion (Figure 4.17(e-f)). The resulting FWI vertical-velocity model (referred to as RTT+FWI model in the following) shows the following structures (Figure 4.17b) : a shallow reflector at 600 *m* in depth, some gas layers between 1500 *m* and 2500 *m* in depth, and a deep reflector below the reservoir level, which is more clearly imaged in the anisotropic reverse-time migration images (Figures 4.18 and 4.19).

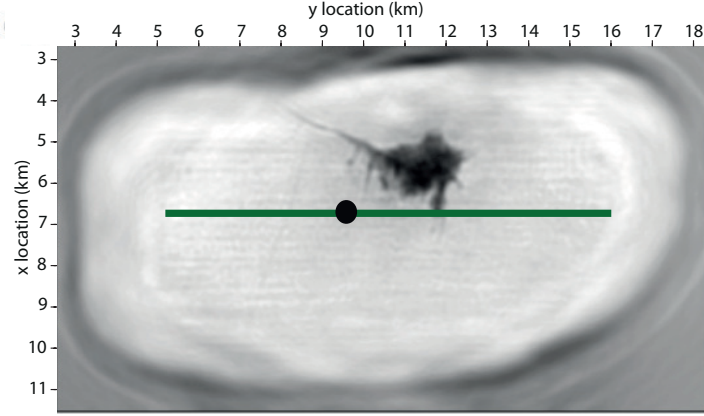


FIGURE 4.15 – The Valhall target. Horizontal slice at a depth of 1000 *m* across the gas cloud extracted from the 3-D FWI model of Sirgue *et al.* (2010). The green line matches the position of cable 21, which is out of the gas cloud. The black circle gives the position of the sonic log shown in this study.

4.0.6.1 Stereotomographic inversion

Multiscale isotropic stereotomography is applied to the real Valhall case study. Since the stereotomography code which is used in this study is isotropic, we expect some footprint of anisotropy in the velocity model building as we shall see in the following. For our application, we consider two data-sets (Figure 4.20). The first was preprocessed by the PGS company for multiple attenuation, and it contains mainly reflections, with a maximum offset of 5200 *m* (Figure 4.20b). The second data-set contains the full wavefield with a maximum offset of 13200 *m*, and it was used for full waveform inversion by Prieux *et al.* (2011) (Figure 4.20a). For stereotomography, we pick the reflection traveltimes and slopes on the first data-set and the early-arriving phases associated with diving waves on the second data-set. We sequentially invert scattering angles greater than 130°, 70°, 40° and 0° and progressively decrease the vertical spacing between the B-splines knots from 500 *m* to 125 *m*.

Stereotomographic models at successive multiscale steps are shown in Figure 4.21 without and with migrated facets. A vertical log from these models at a distance of 9500 *m* is shown in Figure 4.22. The stereotomography model obtained after the step 1 ($\theta > 130^\circ$) is equivalent to a velocity model that would have been obtained by first-arrival traveltime tomography. The migrated facets are located at a maximum depth of 2000 *m*, which leads to first-arrival ray coverage close to that shown in Figure 4.16(c-d). As expected, horizontal velocities rather than the normal moveout or vertical velocities are reconstructed down to 1500 *m* in depth from the wide-aperture picks (Figure 4.22a). This shows that stereotomography succeeds in

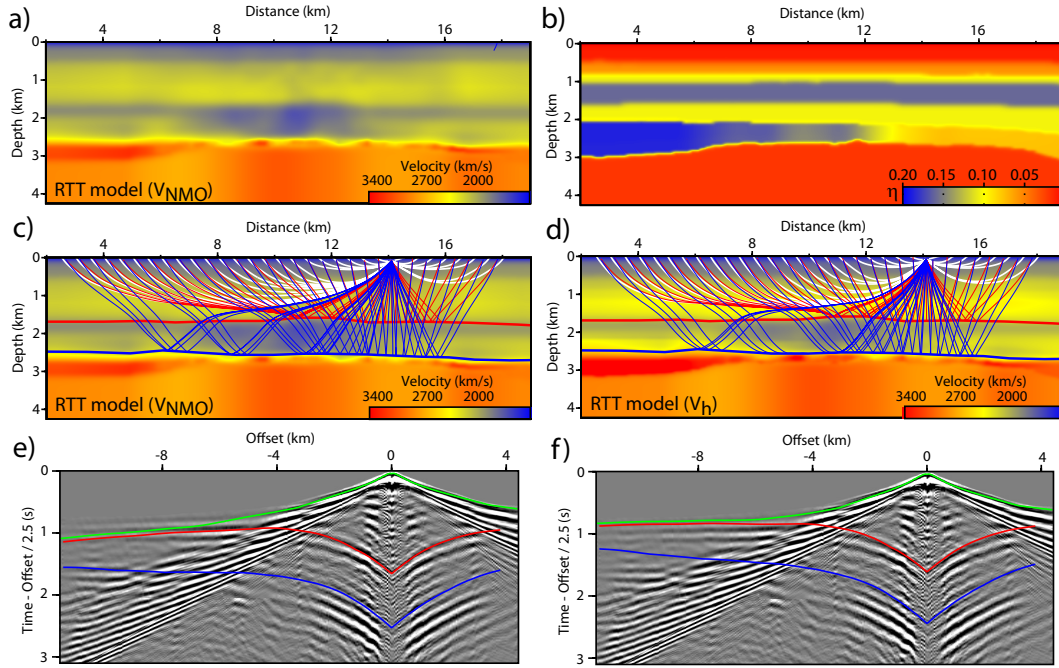


FIGURE 4.16 – Footprint of anisotropy in Valhall. (a) NMO velocity model developed by anisotropic reflection-traveltime tomography (courtesy of BP). (b) Model of η , which shows the strength of the anisotropy (courtesy of BP). (c) Rays tracing in the NMO velocity model for the first arrival (white rays), and the reflections from the top of the gas (red) and the reservoir (blue). Top of the gas and the reservoir are delineated by red and blue solid lines, respectively. Bottom panel : receiver gather with superimposed traveltimes computed in the RTT model for these 3 phases. (d) As (c) for the horizontal velocity model. Note the delay of the first-arrival traveltimes computed in the NMO velocity model.

inverting early-arrival traveltimes, a key result of this study. The poor reconstruction of the low-velocity zone associated with the gas layers and the high-velocity contrast at the reservoir level highlights insufficiently long offset coverage to sample these targets with first-arrival rays.

As stereotomography proceeds towards short apertures, the deep part of the medium is updated, as shown by the position in depth of the migrated facets in Figures 4.21(e-h). The stereotomographic inversion becomes unstable at the fourth step (Figures 4.21d and 4.22d) and we will consider the stereotomographic model at stage 3 as the final stereotomographic model (Figure 4.21c). This instability might be partly due to anisotropy, which makes the wide-aperture and the short-aperture picks inconsistent when isotropic imaging is performed. The final stereotomographic model succeeds in the reconstruction of the large-scale structure of the gas layers, as well as the increase in the velocities at the reservoir level.

4.0.6.2 Full waveform inversion

Anisotropic full waveform inversion is applied following the same experimental set-up as in Prioux *et al.* (2011). Full waveform inversion is performed with a discontinuous Galerkin modeling engine, and five groups of frequencies between 3.5 Hz and 6.7 Hz are successively inverted,

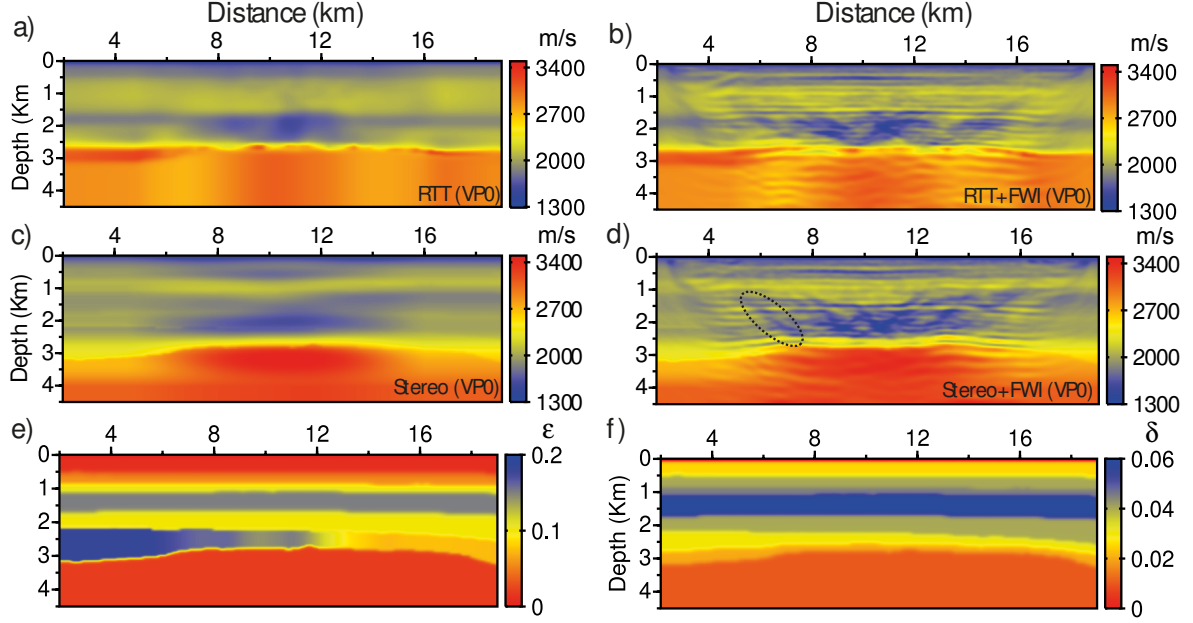


FIGURE 4.17 – Initial and full waveform inversion models. (a) Initial V_{P0} model developed by reflection-traveltime tomography (courtesy of BP). (b) Final V_{P0} FWI model inferred from the initial V_{P0} model shown in (a). (c) Stereotomography model after conversion in vertical velocity. (d) Same as (b) for the initial V_{P0} model shown in (c). The ellipse highlights dipping artifacts (see text for details). (e-f) Background models of ϵ (e) and δ (f).

without time damping. Free-surface multiples are involved in the inversion. The density model is inferred from the initial velocity model, using the Gardner law, and it is kept fixed during the inversion. A homogeneous background attenuation model below the sea bottom is used for seismic modeling, with the quality factor Q_P equal to 150. Before applying anisotropic full waveform inversion for vertical velocity, we need to transform the isotropic stereotomography velocity model into a vertical velocity model. As above-mentioned, the stereotomography model mostly represents horizontal velocities in the upper part of the subsurface because the early steps of the multiscale stereotomography were driven by the early arrivals. We assume that the full stereotomography model represents horizontal velocity, and apply the transformation : $V_{P0}^{stereo} = V^{stereo} / \sqrt{1 + 2\epsilon}$ using the ϵ model of Figure 4.17e, where V^{stereo} and V_{P0}^{stereo} denote the isotropic stereotomography model (Figure 4.21c) and the vertical-velocity stereotomographic model (Figure 4.17c), respectively. Indeed, this is an approximation because the deep part of the isotropic stereotomographic model is supposed to represent NMO velocities rather than horizontal velocities, this part of the subsurface being sampled by short-spread reflections only.

The final FWI model obtained (referred to as stereotomography + FWI model in the following) is shown in Figure 4.17d. The shallow reflector at 0.6 km in depth has been successfully reconstructed, as well as a sharp low-velocity interface, which delineates the top of the gas layers at 1.5 km in depth. The top of the reservoir at 2.5 km in depth is also consistent in the FWI models inferred from the reflection-traveltime-tomography model and the stereotomography model. The deep reflector at 3.5 km in depth is also imaged successfully. The FWI model shows the build-up of the gas layers between 1500 m and 2500 m in depth, although these layers are polluted by dipping artifacts (Figure 4.17d, ellipse). These artifacts, also visible to a lesser

LA STÉRÉOTOMOGRAPHIE COMME OUTIL DE CONSTRUCTION D'UN MODÈLE INITIAL POUR L'INVERSION DES FORMES D'ONDES

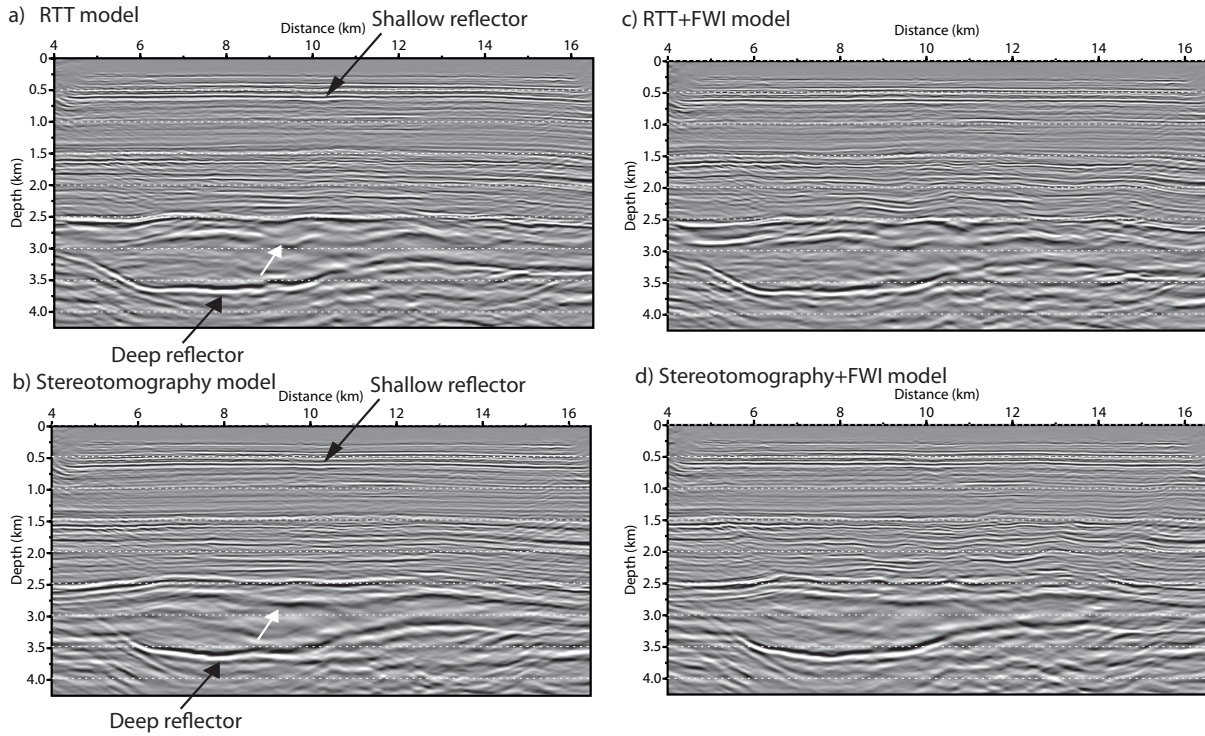


FIGURE 4.18 – Reverse time migrated images computed in the RTT (a), stereotomographic (b), RTT+FWI (c), and stereotomographic + FWI (d) models. The black arrow points on a shallow reflector at 0.6 km in depth and a deep reflector mentioned in the text. The white arrow points on a reflector below the reservoir level, which has been improved in the reverse-time-migration image computed in the stereotomography model.

extent in the RTT+FWI model (Figure 4.17b), are probably a footprint of the anisotropy.

Comparison between a band-pass filtered sonic log of vertical velocity and the corresponding log of the stereotomography+FWI model shows a good agreement down to 2 km in depth, which has been improved relative to the one obtained with the RTT+FWI model (Figure 4.23). Velocities in the gas layers between 2 km and 2.5 km in depth are underestimated in the RTT+FWI and stereotomography+FWI models, this velocity underestimation being more significant in this latter model. This might result again from the crude isotropic-to-*VTI* conversion of the stereotomography model. A noticeable feature is the velocity discontinuity built by full waveform inversion at the reservoir level at 2.5 km in depth, which is consistent with the high velocity contrast of the sonic log at this depth (Figure 4.23b, ellipse). Although this velocity contrast is underestimated in the FWI log, it was absent in the stereotomography model. As full waveform inversion succeeds in reconstructing this velocity contrast, we conclude that the large-scale stereotomography model is sufficiently-kinematically accurate to image the top of the reservoir by full waveform inversion. This is further supported by the reverse-time-migration image computed in the stereotomography model, which has imaged the top of the reservoir at its correct position, and by the focusing of the deep reflector at 3.5 km in depth (Figure 4.18b).

The anisotropic reverse-time-migration images computed in the *VTI* stereotomography and in the reflection-traveltime-tomography (RTT) models show reflectors at similar depths (Fi-

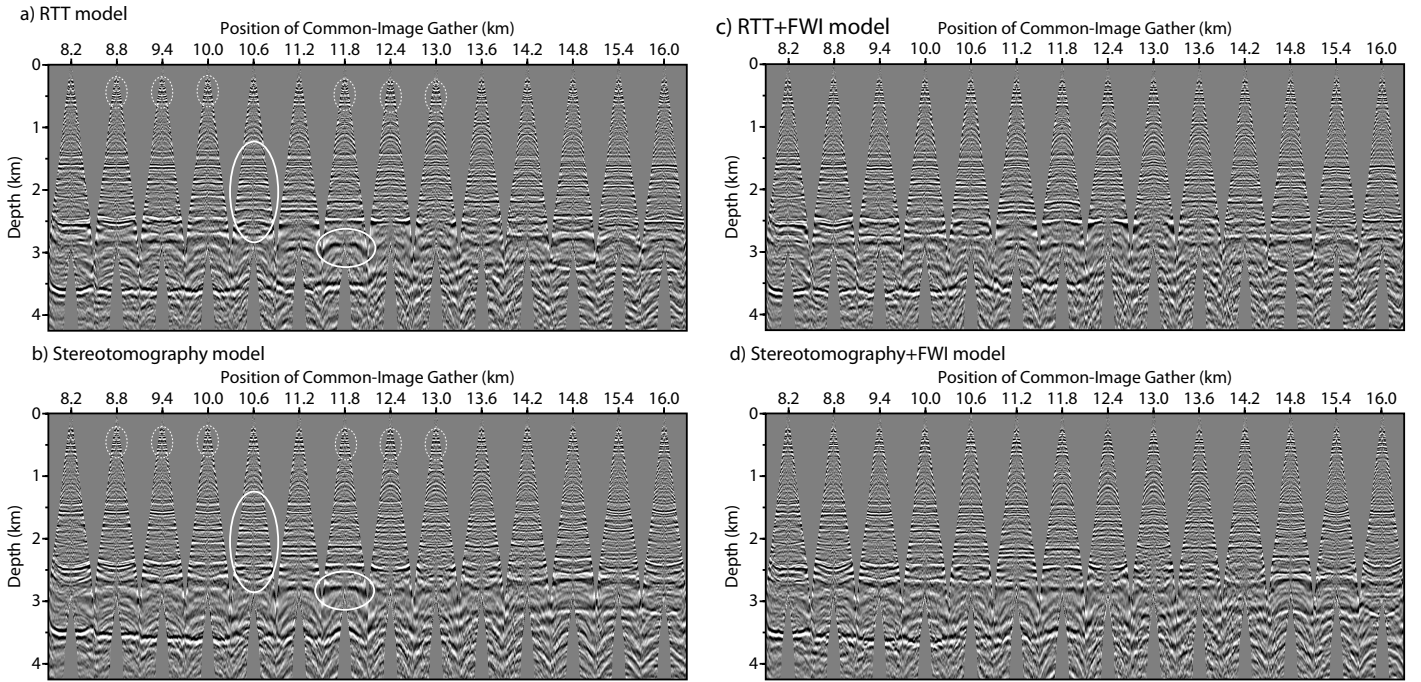


FIGURE 4.19 – Common-image gathers in the depth-offset domain corresponding to the reverse-time-migration (RTM) images of Figure 4.18. The ellipses highlight some reflectors in the shallow part, in the gas and below the reservoir, which are flatter in the stereotomography+RTM image.

figures 4.18(a-b) and 4.19(a-b)). It is worth noting that the reverse-time-migration image computed in the stereotomography model provides a more continuous image of the reflector below the top of the reservoir at around 2.75 km in depth (Figure 4.18(a-b), white arrow). This is supported by the common image gathers computed in the stereotomography model which are locally flatter than the ones computed in the RTT model between 1.5 and 3 km in depth (Figure 4.19(a-b), solid white ellipses). A close inspection of the common image gathers in the shallow part of the target (around 0.5 km in depth) also shows that the stereotomography common image gathers are flatter than the reflection-traveltime-tomography ones, these latter being generally smiling (Figure 4.19(a-b), dashed white ellipses). This improvement probably results from the use of early-arrival traveltimes in stereotomography, which may help to compensate for a deficit of reflection picks from shallow reflectors. Full waveform inversion generally fails to improve the reverse-time-migration images (Figure 4.18(c-d)). This supports that, for this case study where the subsurface does not contain structures of complex geometry as salt bodies, a smooth model is more suitable for reverse time migration, because it honors the scale separation between the background model and the reflectivity image underlying conventional seismic reflection imaging workflow. The footprint of the artificial dipping structures built by FWI (Figure 4.17d) is clearly visible in the reverse-time-migration images computed in the FWI models (Figure 4.18d), and takes the form of suspicious undulations of the reflectors in the gas layers. Of note, these artifacts are also visible in the reverse-time-migration image computed in the RTT+FWI model, and might result from inaccurate accounting of anisotropy during FWI

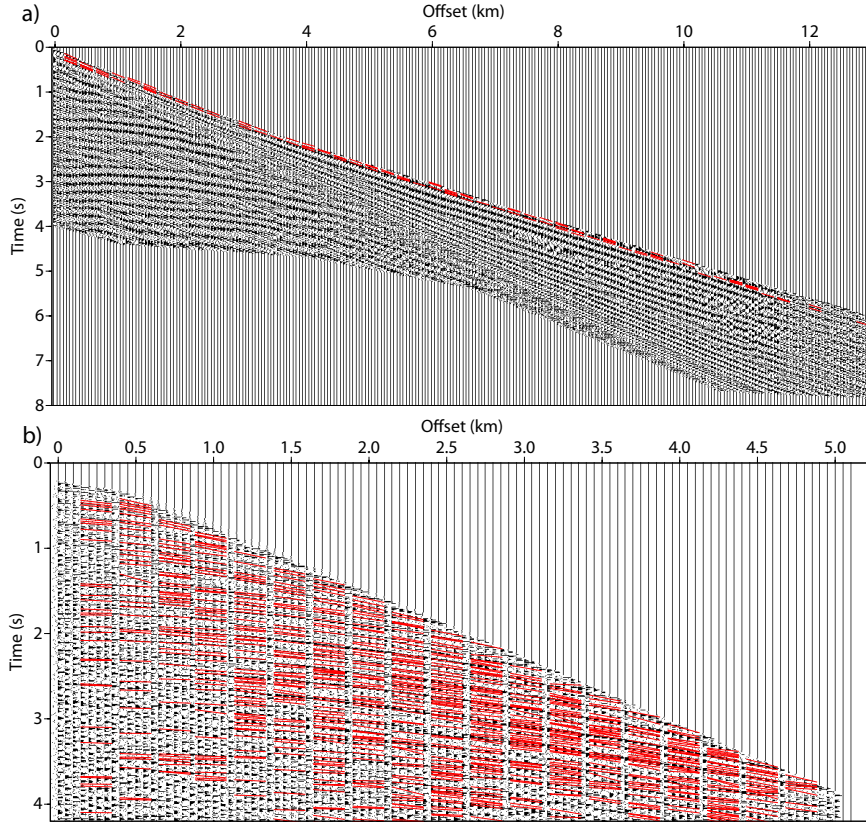


FIGURE 4.20 – Real data case study - (a) Example of receiver gather used to pick early-arrival traveltimes. (b) Example of shot gather used to pick reflections. Slopes of locally-coherent events are superimposed on each gathers. Data are shown after mute and automatic gain control.

(Prioux *et al.*, 2011).

Synthetic seismograms computed in the anisotropic reflection-traveltime-tomography and stereotomography models are shown in Figures 4.24 and 4.25. The match between recorded and modeled seismograms has been clearly improved after FWI, for both the reflection-traveltime-tomography and the stereotomography models. However, the stereotomography+FWI model allows us to achieve a better match of the diving waves relative to the RTT+FWI model, while both FWI models lead to a nearly equivalent match of the reflections. This is an additional evidence of the ability of stereotomography to perform joint inversion of refraction and reflection traveltimes, a key property to build accurate velocity model for FWI from wide-aperture data sets.

4.0.7 Conclusion

We have presented here a non-conventional application of stereotomography to long-offset data, where refraction and reflection traveltimes are jointly inverted during a multiscale inversion, which proceeds from the wide apertures to the shorter ones. The stereotomography is assessed as a semi-automatic tool to build an initial model for FWI. In this context, the joint inversion of refraction and reflection traveltimes allows us to use kinematic information during

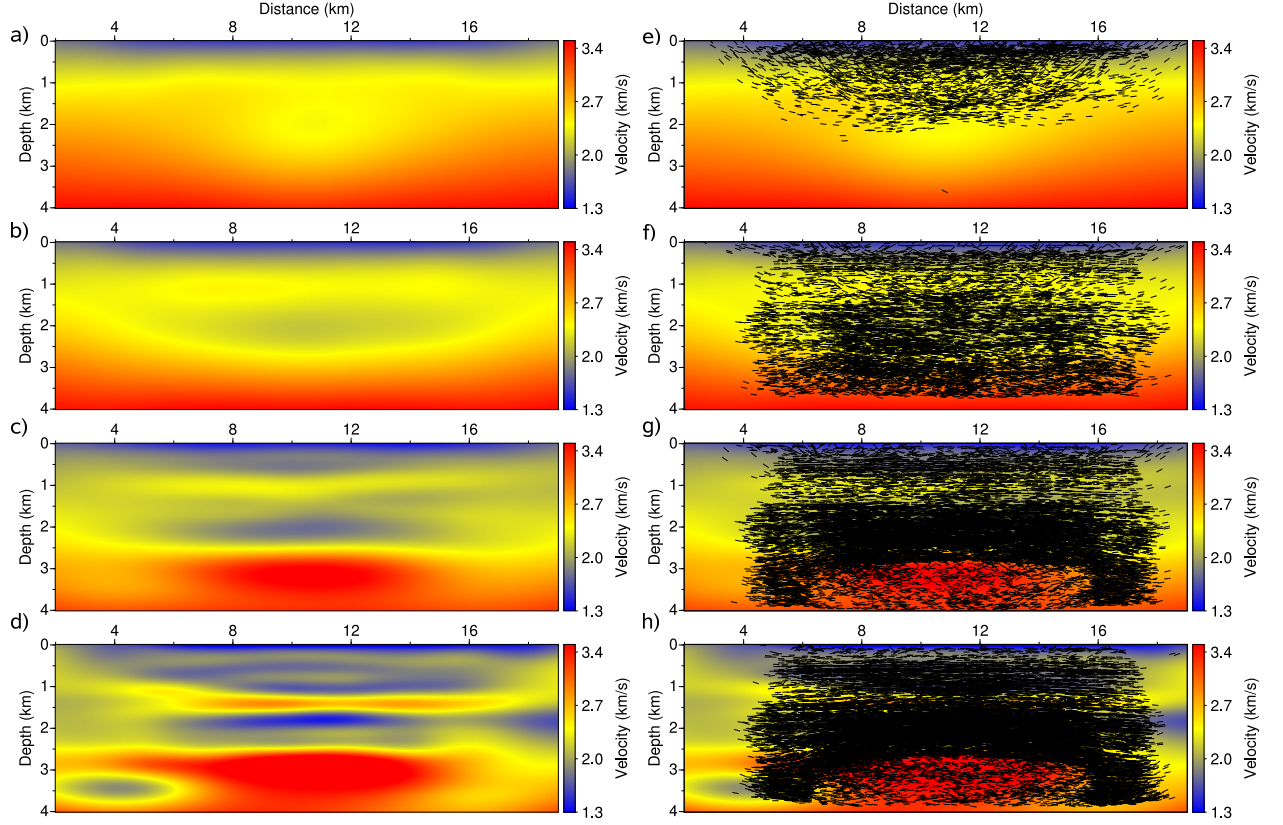


FIGURE 4.21 – Real data case study - Stereotomographic models, plotted without (a-d) and with (e-h) migrated facets, at intermediate stages of the hierarchical stereotomography. (a, e) $\theta > 130^\circ$, (b, f) $\theta > 70^\circ$, (c, g) $\theta > 40^\circ$, (d, h) $\theta > 0^\circ$.

the initial model building, which is as close as possible to that used during FWI. This implies the picking of as many single-scattered events as possible in the full waveform. We first show that joint refraction and reflection stereotomography leads to a velocity model with a much higher resolution than first-arrival traveltime tomography. Second, compared to reflection traveltime tomography, we show that use of refraction traveltimes can be helpful to compensate for a deficit of reflection ray-path illumination by undershooting, in particular in areas where the velocity gently decreases with depth, or below a natural screen such as gas layers. To handle the additional non-linearity introduced by the long-offset refracted picks, we propose a hierarchical formulation of stereotomography, which injects progressively shorter aperture angles into the inversion. We present an application of joint refraction and reflection stereotomography and FWI to a wide-aperture data set from the Valhall field. Clearly, the significant anisotropy of the Valhall data-set has impacted on the results of the joint refraction and reflection stereotomography, that was applied in the isotropic approximation. However, an approximate correction of the anisotropic effects in the stereotomography model allows us to perform reliable anisotropic full waveform inversion and reverse time migration using the stereotomography model as initial or background model. The good match between recorded and modeled seismograms for both diving waves and reflected waves is an additional evidence of the relevance of the stereotomography model as initial model for full waveform inversion. Taking into account for anisotropy will be a key to augment the value of joint refraction and reflection stereotomography, when

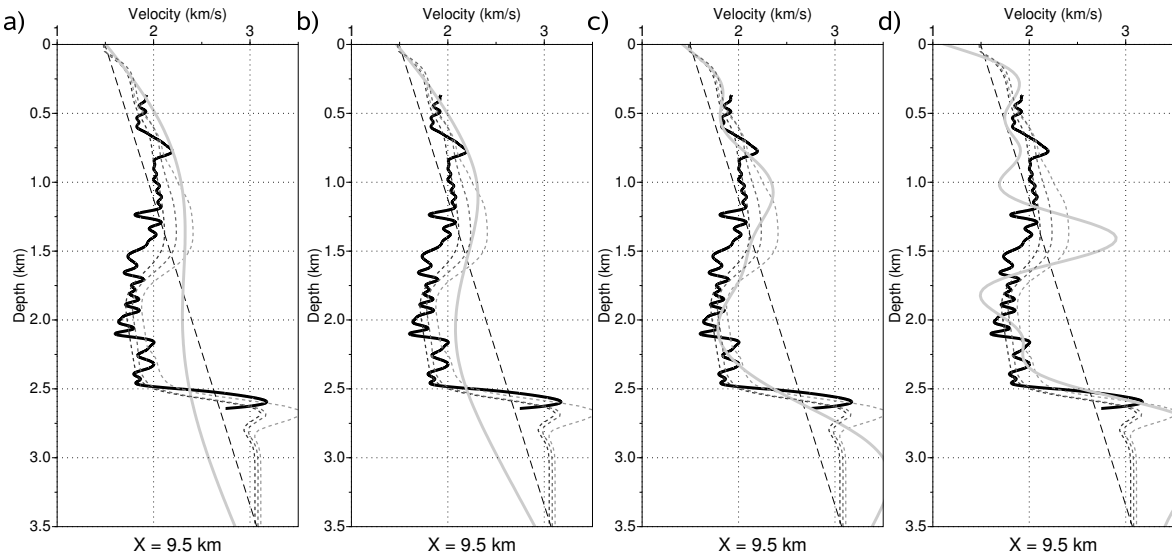


FIGURE 4.22 – Real data case study - Comparison between sonic and isotropic stereotomographic logs at the different stages of the multiscale stereotomography. The black thick line represents a bandpass filtered sonic profile provided by BP. In (a-d), the thick gray lines are extracted from the stereotomographic models shown in Figure 4.21(a-d). The long dashed black curve represent the starting model used by the stereotomography. The stereotomographic logs are compared with profiles extracted from the vertical velocity (dark dash gray), the NMO velocity (dash gray) and the horizontal velocity (light dash gray) models built by anisotropic reflection-traveltime tomography (Figures 4.16(a-b) and 4.17a). Note how the isotropic stereotomography log in (a) matches the large-scale variation of the horizontal velocity in the first 1.5 km in depth, where diving waves propagate.

applied to wide-aperture and wide-azimuth data, and this will be the aim of our future work.

Acknowledgments

This study was funded by the SEISCOPE consortium <http://seiscope.oca.eu>, sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The linear systems were solved with the MUMPS package, available on <http://graal.ens-lyon.fr/MUMPS/index.html> and <http://mumps.enseiht.fr>. The mesh generation was performed with help of TRIANGLE, available on <http://www.cs.cmu.edu/~quake/triangle.html>. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and of CINES/IDRIS under the allocation 2010- [project gao2280] made by GENCI (Grand Equipement National de Calcul Intensif), and we gratefully acknowledge both of these facilities and the support of their staff. We would like to thank H. Chauris (Mines ParisTech) for providing us the stereotomographic code and the 2D ray+Born migration code. We also thank BP Norge AS and their Valhall partner Hess Norge AS for allowing access to data to the Valhall data set as well as the data set preprocessed by PGS for migration and the well-log velocities. We would like to thank the

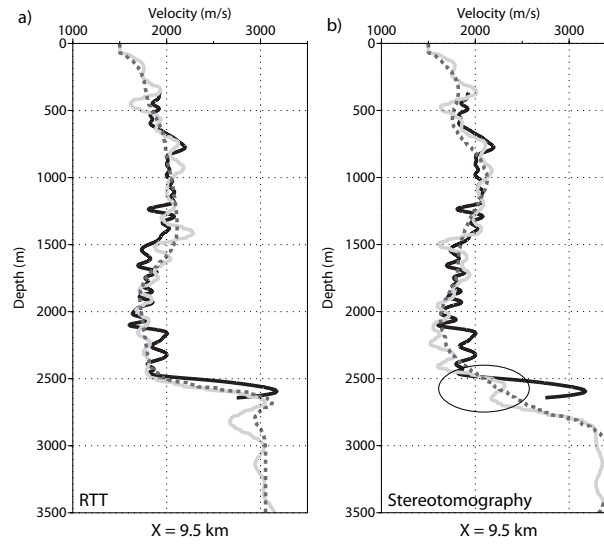


FIGURE 4.23 – Real data case study - Comparison between sonic, traveltime (& slope) tomography and FWI logs. In (a-b), the sonic, the traveltime (& slope) tomographic and the FWI logs are plotted with black, dashed gray and solid gray lines, respectively. The tomographic logs are extracted from the RTT (a) and stereotomography (b) V_{P0} models. The ellipse in (b) highlights the high-velocity contrast that was built by full waveform inversion on top of the reservoir, starting from the stereotomography model.

associate editor Ian Jones and three anonymous reviewers for their comments as well as Richard Wombell (CGG-Veritas) for careful reading the manuscript.

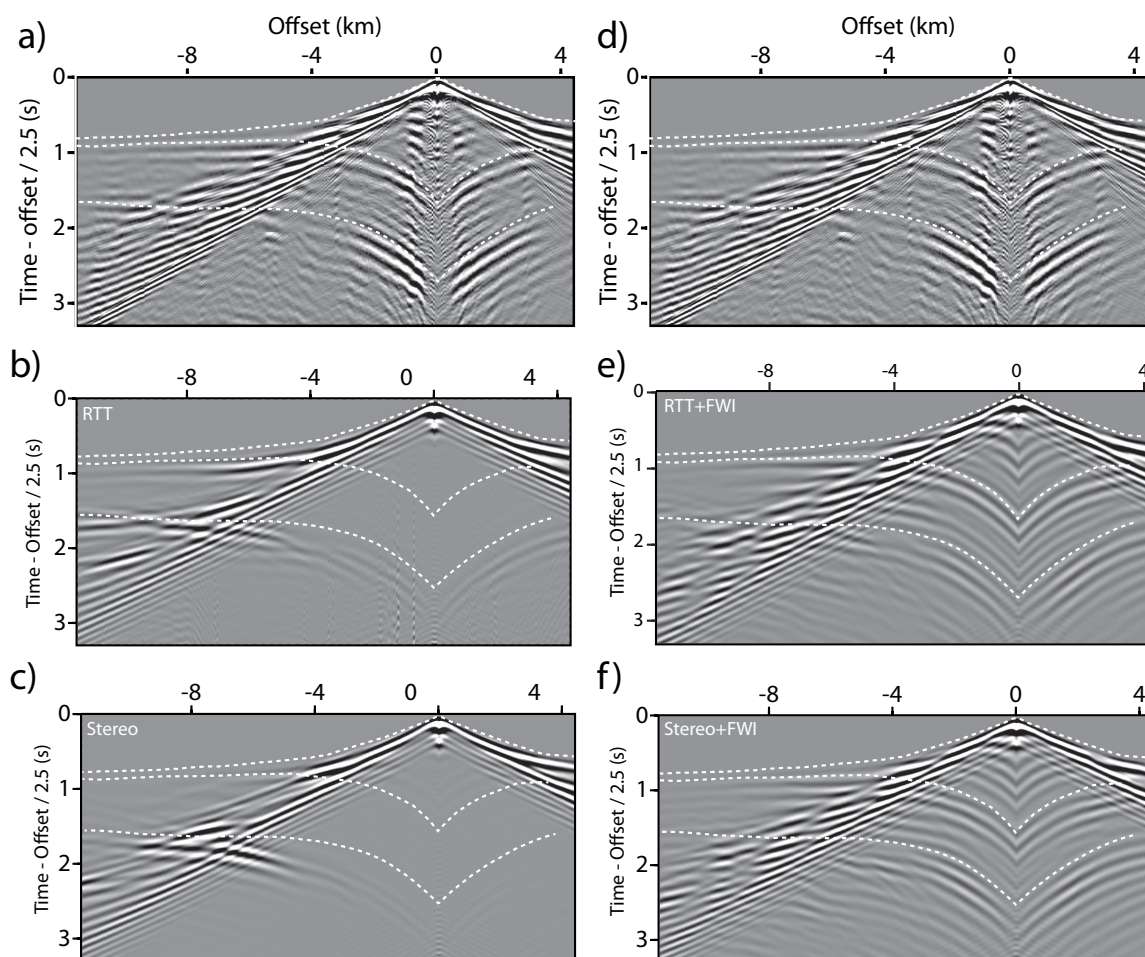


FIGURE 4.24 – (a, d) Recorded receiver gather. (b,c)-(e-f) Corresponding anisotropic synthetic receiver gather computed in the RTT (b), stereotomography (c), RTT+FWI (e) and stereotomography+FWI (f) models. The dash lines represent picked first-arrival traveltimes and reflection traveltimes from the top and bottom of the gas layers computed in the anisotropic reflection-traveltime tomography model.

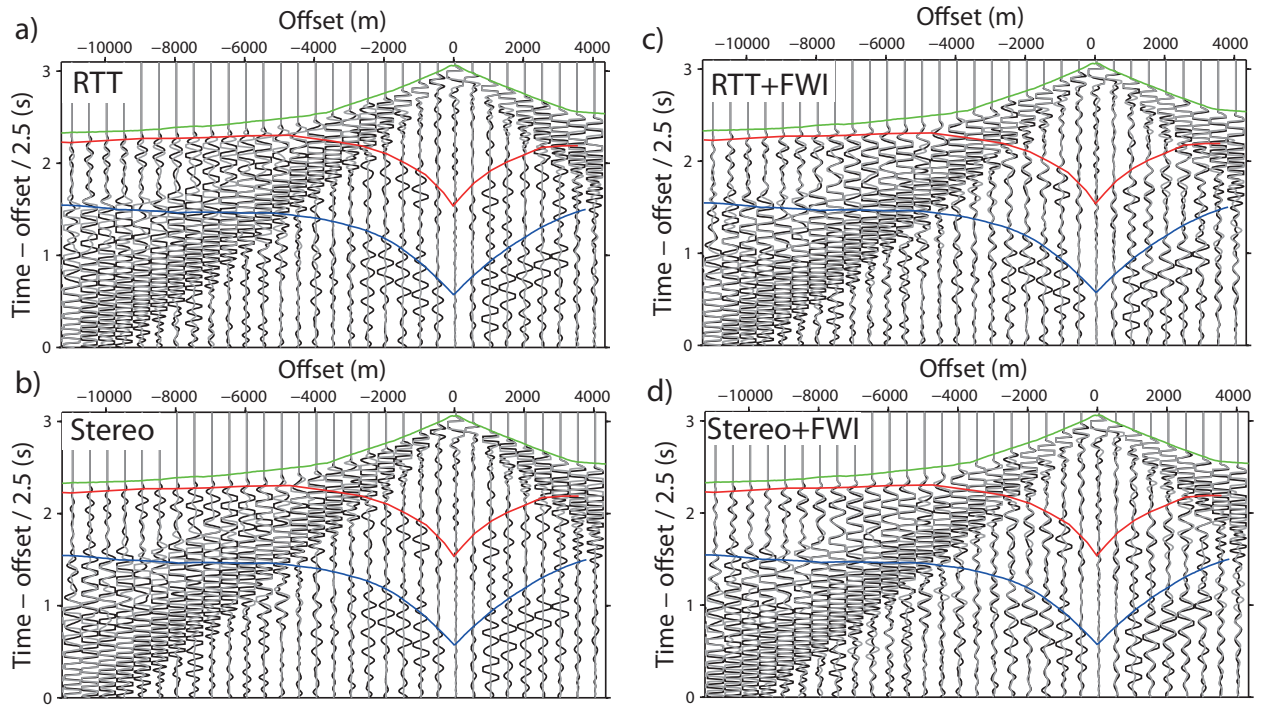


FIGURE 4.25 – Direct comparison between recorded (black) and computed (gray) seismograms shown in Figure 4.24. The yellow curves denote the first-arrival traveltimes computed in the model within which the synthetic seismograms were computed. The green, red and blue lines have the same meaning as the dash white lines in Figure 4.24.

4.0.8 Appendix A : Regularization of multi-data / multi-parameter inversions

In this Appendix, we review the inversion algorithm that was developed by Billette & Lambaré (1998); Billette *et al.* (2003); Lambaré *et al.* (2004a) to handle multiple classes of data and parameters. We consider a new set of normalized model parameters \mathbf{p} for inversion, which are related to the physical parameters \mathbf{m} by a diagonal weighting matrix \mathbf{W}_m

$$\mathbf{m} = \mathbf{W}_m \mathbf{p}. \quad (4.3)$$

The aim of the normalization operator \mathbf{W}_m , the expression of which is given at the end of this section, is to better balance the weight of each parameter class in the inversion. Eq. 4.3 leads to :

$$\mathbf{J}_p = \mathbf{J}_m \mathbf{W}_m, \quad (4.4)$$

where \mathbf{J}_p and \mathbf{J}_m denote the Fréchet derivative or sensitivity matrices for the \mathbf{p} and \mathbf{m} parametrizations, respectively. The expression of the Fréchet derivatives, which are computed by means of paraxial ray tracing, are provided in Billette & Lambaré (1998, their Appendix 1).

We minimize the least-squares misfit function with respect to \mathbf{p} subject to smoothing regularization on the velocity perturbations

$$\mathcal{C}(\mathbf{p}) = \frac{1}{2} \left[(\mathbf{d}^m - \mathbf{d}^r)^T \mathbf{W}_d (\mathbf{d}^m - \mathbf{d}^r) + \lambda (\mathbf{p} - \mathbf{p}_{prior})^T \mathbf{R}^T \mathbf{L}^T \mathbf{L} \mathbf{R} (\mathbf{p} - \mathbf{p}_{prior}) \right], \quad (4.5)$$

where the diagonal matrix $\mathbf{W}_d = \text{diag} \left\{ 1/\sigma_S^2, 1/\sigma_R^2, 1/\sigma_{TSR}^2, 1/\sigma_{pS}^2, 1/\sigma_{pR}^2 \right\}$ contains the standard deviations of each class of data, \mathbf{L} is a Laplacian operator applied to the normalized velocity perturbations, \mathbf{R} is a restriction operator to the velocity perturbations, \mathbf{d}^r and \mathbf{d}^m are the recorded and modeled stereotomographic datasets, \mathbf{p}_{prior} is the prior model. In this study, we use $\sigma_{S_x} = \sigma_{S_z} = \sigma_{R_x} = \sigma_{R_z} = 5 \text{ m}$, $\sigma_{TSR} = 5 \text{ ms}$, $\sigma_{pS} = \sigma_{pR} = 0.01 \text{ s/m}$. The model perturbations $\Delta \mathbf{p}$ satisfy the system of normal equations

$$\left(\mathbf{J}_p^{(k)T} \mathbf{W}_d \mathbf{J}_p^{(k)} + \epsilon \mathbf{R}^T \mathbf{L}^T \mathbf{L} \mathbf{R} \right) \Delta \mathbf{p}^{(k)} = \mathbf{J}_p^{(k)T} \mathbf{W}_d \Delta \mathbf{d} + \mathbf{L} \left(\mathbf{p}^{(k)} - \mathbf{p}_{prior} \right). \quad (4.6)$$

In this study, we use $\mathbf{p}_{prior} = \mathbf{p}^{(k)}$, for all k , which allows the second term of the right-hand side of Eq. 4.6 to be dropped out :

$$\left(\mathbf{J}_p^{(k)T} \mathbf{W}_d \mathbf{J}_p^{(k)} + \epsilon \mathbf{R}^T \mathbf{L}^T \mathbf{L} \mathbf{R} \right) \Delta \mathbf{p}^{(k)} = \mathbf{J}_p^{(k)T} \mathbf{W}_d \Delta \mathbf{d}, \quad (4.7)$$

Substituting Eqs. 4.3 and 4.4 in Eq. 4.7 gives the expression of the physical model perturbations

$$\Delta \mathbf{m}^{(k)} = \mathbf{W}_m \left(\mathbf{W}_m^T \mathbf{J}_m^{(k)T} \mathbf{W}_d \mathbf{J}_m^{(k)} \mathbf{W}_m + \epsilon \mathbf{R}^T \mathbf{L}^T \mathbf{L} \mathbf{R} \right)^{-1} \mathbf{W}_m^T \mathbf{J}_m^{(k)T} \mathbf{W}_d \Delta \mathbf{d}. \quad (4.8)$$

It can be seen that the model reparametrization, Eq. 4.3, is equivalent to the normalizing of each column of the Fréchet derivative matrix by the corresponding diagonal element of \mathbf{W}_m .

For the diagonal coefficients w_{ll} of \mathbf{W}_m , we use :

$$w_{ll} = 1 / \left\{ \sqrt{\sum_{i=1}^{7 \times N} \left(\frac{\partial d_i}{\partial m_l} \frac{1}{\sigma_i} \right)^2} / H + \kappa \right\} \quad (4.9)$$

where $H = \text{Max}_l \left\{ \sqrt{\sum_{i=1}^{7 \times N} \left(\frac{\partial d_i}{\partial m_l} \frac{1}{\sigma_i} \right)^2} \right\}$, κ is a damping factors that prevents division by zero (Wang, 1993). Of note, the number of rows of the Fréchet derivative matrix is $7 \times N$ which corresponds to four coordinates for the shot and receiver positions, one two-way traveltime, and two slopes per event. The diagonal matrix \mathbf{W}_m provides a reliable weighting of each column of \mathbf{J} , such that each parameter class has a similar contribution in the Fréchet derivatives matrix. Moreover, it attempts to reduce the effects of the non uniform ray coverage of the velocity parameters, by increasing the weights of the poorly-illuminated parameters at the expense of the well-illuminated ones.

Each coefficient of column l of the Fréchet derivative matrix is divided by the least-squares norm of column i of \mathbf{J} normalized by the maximum of the coefficients, such that the coefficients vary between 0 and 1.

The damping coefficient λ which balances the relative contribution of the least-squares norm of the misfit vector and the Laplacian regularization, is defined as :

$$\lambda = p2d \frac{\text{Tr} \left(\mathbf{W}_m^T \mathbf{J}_m^{(k)T} \mathbf{W}_d \mathbf{J}_m^{(k)} \mathbf{W}_m \right)}{\text{Tr} \left(\mathbf{R}^T \mathbf{L}^T \mathbf{L} \mathbf{R} \right)} \quad (4.10)$$

where we use $p2d = 10 \%$ and Tr denotes the trace of a matrix (Wang, 1993).

In practice, we build the augmented tomographic system for \mathbf{p} :

$$\begin{pmatrix} \Delta \mathbf{d} / \sigma^2 \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{J}_m \mathbf{W}_m / \sigma^2 \\ \lambda \mathbf{R} \mathbf{L} \end{pmatrix} \Delta \mathbf{p} \quad (4.11)$$

and this system is solved with the LSQR conjugate gradient method. The physical model perturbations $\Delta \mathbf{m}$ are subsequently inferred from $\Delta \mathbf{p}$ using Eq. 4.3.

4.0.9 Conclusion de l'article

La représentation des données stéréotomographiques sur les sismogrammes a été déterminante dans ma phase de recherche, car cela m'a permis de faire le lien entre les données et les angles d'ouverture calculés au niveau du point diffractant. J'ai de plus pu faire la relation entre la nature des phases inversées que je pouvais visualiser, et les structures imagées par stéréotomographie. J'ai ainsi pu discriminer les arrivées enregistrées après la première arrivée et avant l'onde directe qui aboutissaient à des instabilités, ou encore les pointés trop courts offsets inutiles en raison de l'indétermination vitesse-profondeur.

Pour étendre la stéréotomographie à l'utilisation des données grands angles, nous nous sommes inspirés de la stratégie des amortissements temporels en inversion des formes d'ondes, en proposant l'introduction séquentielle de pointés stéréotomographiques dont l'angle d'ouverture au niveau du point diffractant est de plus en plus petit au cours de l'inversion. Cela a permis de minimiser la non-linéarité de l'inversion en reconstruisant hiérarchiquement les nombres d'ondes du modèle. L'application de cette procédure au champ de Valhall a montré des résultats satisfaisants, bien que la résolution soit inférieure à ce qui avait été obtenu par tomographie en réflexion. Ces résultats pourraient être améliorés en introduisant l'anisotropie dans l'outil de modélisation de la stéréotomographie (Nag *et al.*, 2010). L'extension de la méthode à des dispositifs lâches d'acquisition serait très intéressante, notamment pour des jeux

LA STÉRÉOTOMOGRAPHIE COMME OUTIL DE CONSTRUCTION D'UN MODÈLE INITIAL POUR L'INVERSION DES FORMES D'ONDES

de données académiques, ou pour le pointé suivant la direction *cross-line* des acquisitions 3D, souvent moins dense en récepteurs que la direction *in-line* (Alerini *et al.*, 2009).

Chapitre 5

Inversion visco-acoustique des formes d'ondes - Reconstruction de la vitesse P, de la densité ou l'impédance, et de l'atténuation

Sommaire

5.1	Introduction à l'article	175
5.1.1	La vitesse des ondes P	176
5.1.2	Masse volumique et impédance	176
5.1.3	L'atténuation	177
5.1.4	Position du problème multiparamètre	180
5.2	Résumé	182
5.3	Introduction	183
5.4	Inversion des formes d'ondes visco-acoustique	184
5.5	Reconstruction de la densité ou de l'impédance	187
5.5.1	Paramétrisations et diagrammes de rayonnement	187
5.5.2	Cas synthétique de Valhall	190
5.6	Reconstruction de la vitesse, de la densité et du facteur de qualité	
	Q_P	197
5.6.1	Cas synthétique de Valhall	198
5.6.2	Application aux données OBC du champ de Valhall	207
5.7	Conclusion	218
5.8	Conclusion à l'article	218

5.1 Introduction à l'article

Beaucoup d'applications d'inversion des formes d'ondes se sont limitées ces dernières années à la reconstruction de la vitesse des ondes P, qui est le paramètre ayant une empreinte dominante dans les données sismiques (e.g. Ravaut *et al.*, 2004; Plessix, 2009b; Sirgue *et al.*,

2010). L'objectif de ce chapitre est d'évaluer la faisabilité d'une reconstruction multiparamètres dans le cadre de l'approximation visco-acoustique à l'aide de cas d'études synthétique, avant une application aux données réelles du champ de Valhall. Pour celle-ci, les reconstructions de V_P , de la densité et de l'atténuation sont réalisées dans l'approximation visco-acoustique pour des raisons de coûts numériques, bien que l'approximation visco-élastique soit plus appropriée pour décrire le comportement AVO des réflexions PP. L'équation d'onde visco-acoustique isotrope peut être paramétrée par différentes combinaisons de 3 paramètres, parmi lesquels le module d'incompressibilité K , la vitesse de propagation des ondes P (V_P), la densité (ρ), l'impédance (I_P) et le facteur de qualité (Q_P) (l'inverse de l'atténuation). Je commencerai par une brève introduction des principaux paramètres avant de discuter les problèmes théoriques inhérents à la reconstruction de plusieurs classes de paramètres par inversion des formes d'ondes. Je m'intéresserai ensuite plus spécifiquement à la reconstruction des deux paramétrisations en vitesse-densité et en vitesse-impédance sur le cas d'étude synthétique de Valhall. Dans un deuxième temps, j'inclurai l'atténuation dans la paramétrisation de l'inversion, et je testerai différentes stratégies pour reconstruire les trois classes de paramètres. Finalement, j'appliquerai le protocole multiparamètre retenu à l'issue de l'analyse paramétrique sur données synthétiques aux données réelles des lignes 21 et 29 enregistrées sur le champ de Valhall.

5.1.1 La vitesse des ondes P

La vitesse des ondes P est d'autant plus rapide que le module d'incompressibilité K de la roche est grand. Ce module dépend fortement de la lithologie, de la porosité Φ et de la saturation en fluide du milieu. De manière générale, la vitesse est inversement proportionnelle à la porosité ($1/V = A + B\Phi$) (Domenico, 1984). La présence de fluide tend à baisser le module d'incompressibilité moyen de la roche, lequel dépend de la saturation des différentes phases fluides. Lorsque les différentes phases sont présentes à la plus petite des échelles, la moyenne de Reuss donne l'expression du module d'incompressibilité de la phase fluide (K_{fl}) comme $1/K_{fl} = \sum_i S_i/K_i$, où S_i et K_i sont les saturations et les modules d'incompressibilité des différentes phases. Par ailleurs, on peut citer la loi de Gassmann (1951) qui relie à la porosité, à basse fréquence, les modules d'incompressibilité des différentes phases constituant la roche .

5.1.2 Masse volumique et impédance

La masse volumique (ρ , en kg/m^3) d'une roche est une moyenne arithmétique volumétrique sur les masses volumiques des constituants de la roche. La densité d'un corps est le rapport de sa masse volumique à la masse volumique d'un corps pris comme référence. Précisons qu'en anglais, le terme "*density*" dénote la masse volumique. Par la suite, nous utiliserons par abus de langage le terme densité au lieu de masse volumique. La densité est étroitement liée à la porosité ϕ par la relation : $\rho = (1 - \phi)\rho_0 + \phi\rho_{fl}$, où ρ_0 et ρ_{fl} sont la densité des grains et des fluides de pores. Par contre, la vitesse n'est pas linéairement liée à la porosité, et donc à la densité. Par exemple, les fractures et les failles peuvent faire décroître les vitesses V_P et V_S , alors même que les fractures ont une porosité nulle (Mavko *et al.*, 2009). C'est la raison pour laquelle les relations empiriques entre vitesse et densité sont en général plus réalistes dans les roches sous hautes pressions et saturées en fluide, car ces deux effets minimisent l'effet des fractures. Les lois empiriques de Gardner *et al.* (1974) reliant la densité à V_p sont très populaires pour construire un modèle initial de densité. Nous utilisons une loi polynomiale de la forme :

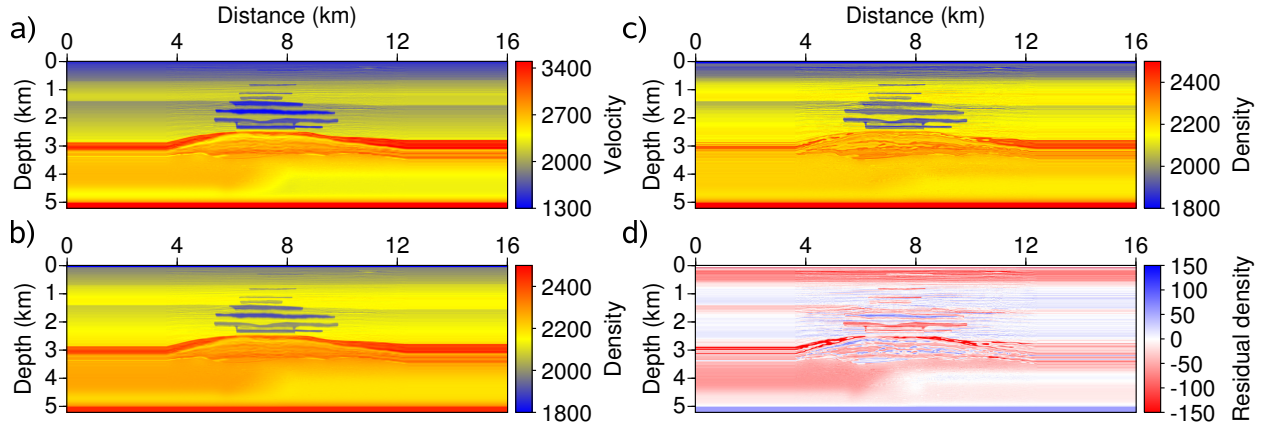


FIGURE 5.1 – (a,c) Modèles synthétique de vitesse (a) et de densité (b) du champ de Valhall. (b) Modèle de densité déduit du modèle de V_P synthétique par une loi de Gardner. (d) Modèle résiduel calculé par la différence de (b) et (c).

$\rho = aV_P^2 + bV_P + c$, où les coefficients a , b et c dépendent du type de roche considéré. Nous avons trouvé dans Mavko *et al.* (2009) les valeurs suivantes correspondant à des argiles, qui nous l'avons vu sont majoritaires dans les sédiments tertiaires de la structure du champ de Valhall : $a = -0.0261$, $b = 0.373$ et $c = 1.458$ pour une gamme de vitesse entre 1.5 et 5 km/s. L'application de cette loi au modèle de vitesse synthétique de Valhall (Figure 5.1a) aboutit au modèle de densité représenté sur la Figure 5.1b, qui est relativement proche du modèle synthétique de densité fourni par BP (Figure 5.1c). Le modèle de densité obtenu en calculant la différence de ces deux modèles (Figure 5.1d) indique que les résidus parmi les plus faibles apparaissent dans les sédiments tertiaires entre 600 m et 2.6 km de profondeur, confortant la validité de la loi empirique pour ce type de sédiments. Les résidus les plus élevés sont situés au toit du réservoir composé de calcaire crayeux. Le fort contraste latéral vers 4 km de profondeur pourrait être un indice de changement lithologique. De plus, le modèle synthétique de densité semble être d'une plus haute résolution que celui de vitesse et plus contrasté au niveau du réservoir, probablement en raison de la nature plus ou moins poreuse et saturée en fluides de ces roches. L'impédance, définie comme le produit de la vitesse et de la densité, est un paramètre clé pour caractériser les interfaces géologiques en vue d'une meilleure interprétation structurale. Les coefficients de réflexion et de transmission aux interfaces dépendent en effet de ce paramètre, qui contrôle de ce fait la quantité d'énergie transmise et réfléchi par une onde incidente sur une interface. Traditionnellement, la reconstruction de l'impédance est préférée à celle de la densité car cette dernière est jugée avoir une influence trop faible sur les données et d'être mal résolue par l'inversion (Tarantola, 1987; Mora, 1987; Forgues & Lambaré, 1997). Nous verrons cependant qu'une paramétrisation en vitesse-densité plutôt qu'en vitesse-impédance pour la reconstruction de ces 3 paramètres peut s'avérer plus avantageuse. Des détails sur l'implémentation des paramétrisations sont donnés dans l'annexe A.

5.1.3 L'atténuation

L'atténuation des ondes sismiques se manifeste sur la forme d'onde par une décroissance de l'amplitude (effet d'absorption) et un élargissement de la largeur du signal, ainsi que par

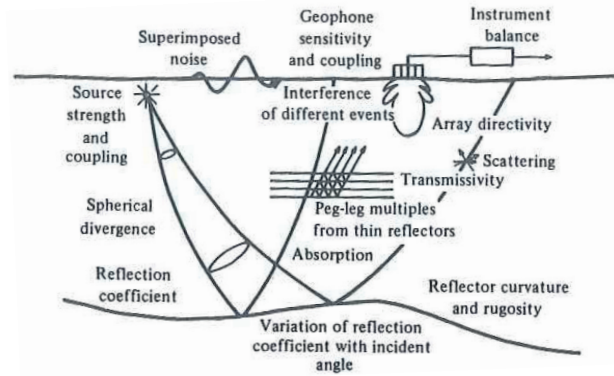


FIGURE 5.2 – Facteurs affectant l'amplitude des ondes (extrait de Sheriff (1975)).

une dispersion fréquentielle qui transporte l'énergie vers l'avant du signal (Toverud & Ursin, 2005). On distingue l'atténuation intrinsèque de l'atténuation extrinsèque. Ces deux phénomènes pouvant être à l'origine des effets sur l'amplitude, il est difficile de distinguer l'une de l'autre. L'atténuation extrinsèque est fonction de la divergence géométrique des ondes, de la géométrie du milieu et de la source, et sont normalement prises en compte dans le calcul du champ d'onde complet. Les diffractions et les multiples internes se produisant lors de la propagation sont inclus dans sa définition. L'atténuation extrinsèque dépend de la fréquence, les ondes basses fréquences se propageant dans des couches dont l'épaisseur est petite par rapport à la longueur d'onde voient un milieu homogène, alors que les plus hautes fréquences se réfléchiront de multiples fois au sein des couches et perdront donc plus d'énergie au cours de la propagation. L'atténuation intrinsèque est associée à la perte d'énergie des ondes par friction se produisant dans des milieux dits 'anélastiques', comme ceux dont les grains ne sont pas élastiquement liés. Elle peut aussi être associée à des mouvements de fluide (ou de gaz) au sein des pores lors du passage de l'onde (Thomsen *et al.*, 1997), et au contenu en argile des pores de la roche (Klimentos & McCann, 1990). Tout comme l'anisotropie, l'atténuation intrinsèque et extrinsèque sont des notions dépendant de la longueur d'onde considérée. Lorsque la taille de l'hétérogénéité est de l'ordre de la longueur d'onde, l'atténuation est sensible aux deux types d'atténuations. L'atténuation que nous voulons extraire des données sismiques est une atténuation effective étant égale à la somme de ces deux origines d'atténuation. Elle est définie par le facteur de qualité Q , dont l'amplitude est inversement proportionnelle à l'atténuation. L'ensemble des effets à l'origine d'une diminution d'amplitude des ondes sont synthétisés dans la Figure 5.2.

La connaissance de l'atténuation permet une meilleure description lithologique, de l'état physique et du degré de saturation des roches, très utile pour mieux caractériser et suivre dans le temps les réservoirs (Toksöz & Johnston (1981) et Wang (2008)). Une des expressions mathématiques du facteur de qualité est donnée par $1/Q(\omega) = -\Delta E/(2\pi E)$, où ΔE est l'énergie perdue à chaque nouveau cycle d'une onde de pulsation ω (Aki & Richards, 2002). Dans notre code de modélisation, nous considérons que l'atténuation est indépendante de la fréquence en raison de la bande étroite de fréquences inversées. Il n'y a pas d'effet de dispersion fréquentielle (variation des vitesses avec la fréquence) dans le cas où Q est indépendant de la fréquence, mais il se produit une déformation du spectre avec la distance se traduisant par un élargissement de la forme d'onde. Cet effet est lié à l'atténuation préférentielle des composantes hautes fréquences propageant un plus grand nombre de longueurs d'ondes pour une distance

donnée que les composantes basses fréquences (Sheriff & Geldart, 1995). L'implémentation de l'atténuation dans un code de simulation en domaine fréquentiel est simple comparativement à une implémentation en domaine temporel, en raison de l'utilisation d'une vitesse complexe $\bar{c} = c(1 - (i \operatorname{sgn}(\omega))/(2Q))$. Le sous-sol est alors décrit par ses propriétés propagatives et diffusives, associées à la partie réelle et imaginaire de la vitesse. Celles-ci peuvent d'ailleurs être reconstruites comme deux paramètres indépendants (Liao & McMechan, 1995; Song & Williamson, 1995; Song *et al.*, 1995; Liao & McMechan, 1996). De nombreux modèles rendent compte de la dépendance de l'atténuation à la fréquence, le plus populaire étant le modèle de Kolski-Futterman (Kolsky, 1956; Futterman, 1962; Toverud & Ursin, 2005) dont la vitesse complexe est donnée par

$$\frac{1}{\bar{c}(\omega)} = \frac{1}{c} \left[\left(1 + \frac{1}{\pi Q} |\log(\omega/\omega_r)| \right) + i \frac{\operatorname{sgn}(\omega)}{2Q} \right]^{-1}, \quad (5.1)$$

où ω_r est la fréquence de référence, prise égale à 1 Hz ou comme la fréquence dominante du signal sismique. Ce modèle a une dépendance fréquentielle linéaire avec le coefficient d'atténuation, tandis que la déviation à la vitesse de phase est prise en compte par un terme variant comme le logarithme de la fréquence. Ce dernier terme est aussi appelé terme de causalité, et permet de diminuer l'ambiguïté existant entre la vitesse et l'atténuation lors de l'inversion des formes d'ondes (Hak & Mulder, 2011). La dépendance de l'atténuation avec la fréquence est attribuée pour les roches poreuses saturées à des mouvements locaux de fluides dans les pores et fissures (Jones, 1986; Sams *et al.*, 1997).

De nombreuses méthodes pour estimer Q à partir de données VSP et d'informations de puits ont été développées. La méthode des rapports spectraux, par exemple, utilise la variation du spectre d'amplitude avec la profondeur pour calculer Q , suivant la formule (Tonn, 1991) : $\ln(|A_2(\omega)|/|A_1(\omega)|) = \alpha - \omega d/(2cQ)$, où $A_1(\omega)$ et $A_2(\omega)$ sont les amplitudes spectrales à différentes profondeurs, d est la distance de propagation entre celles-ci, et α une constante. Des estimations de Q ont aussi été faites à partir de données de sismique de surface avec la méthode des rapports spectraux par Dasgupta & Clark (1998). Par ailleurs, Ribodetti *et al.* (2000a) a démontré que la reconstruction de la vitesse et l'atténuation par inversion des formes d'ondes dans l'approximation asymptotique ray+Born n'était possible que si le réflecteur était illuminé par une onde descendante et ascendante (via un réflecteur sous-jacent par exemple), le Hessien asymptotique devenant dans le cas contraire singulier. Des études de Mulder & Hak (2009); Hak & Mulder (2010) ont montré que, dans le cadre d'acquisitions de sismique réflexion à faibles dépôts, une transformée de Hilbert lie la vitesse et l'atténuation, ce qui empêche la reconstruction indépendante de ces paramètres, de nombreuses combinaisons de modèles de vitesse et d'atténuation pouvant expliquer les données à l'identique. Des applications de reconstruction visco-acoustique par inversion des formes d'ondes ont été réalisées sur cas synthétique par Askan *et al.* (2007) ou Liao & McMechan (1995). Les applications sur données réelles ont principalement été réalisées pour des acquisitions de type "entre-puits" (Song & Williamson, 1995; Pratt *et al.*, 2005; Kamei & Pratt, 2008; Rao & Wang, 2008), et dans une moindre mesure pour des acquisitions de surface (Hicks & Pratt, 2001; Smithyman *et al.*, 2009; Malinowski *et al.*, 2011). Il fut montré par Malinowski *et al.* (2011) que la reconstruction conjointe de la vitesse et de l'atténuation est possible pour des milieux très atténuants à partir des informations d'amplitude et de phase, à la condition que le modèle initial de vitesse soit suffisamment précis. La reconstruction de l'atténuation à basses fréquences peut cependant s'avérer instable en raison de la faible sensibilité du paramètre à ces fréquences, requérant alors un fort lissage et amortissement des perturbations du facteur Q . Sauf mention contraire, nous désignerons dans la suite par atténuation le facteur de qualité.

5.1.4 Position du problème multiparamètre

L'inversion des formes d'ondes est généralement restreinte à la reconstruction de la vitesse des ondes P car ce paramètre présente la plus forte empreinte dans les données. Cependant, les paramètres de second ordre tels que la vitesse de propagation des ondes de cisaillement, la densité et l'atténuation peuvent être essentiels pour mieux caractériser les fluides ou d'autres propriétés pétrophysiques du milieu. De plus, l'omission de paramètres secondaires dans la modélisation peut pénaliser la convergence de l'inversion en raison des erreurs sur l'amplitude des signaux qui en résultent. La cinématique et la signature AVO des données grands offsets peuvent par exemple être très sensibles aux paramètres secondaires tels que les paramètres de Thomsen, l'atténuation et la vitesse de cisaillement. Je passe maintenant en revue les principales difficultés expliquant le faible nombre d'applications d'inversions multiparamètres, et je propose quelques pistes de réflexion pour essayer de les résoudre. On distingue deux grandes catégories de difficultés, associées à la variabilité de l'empreinte des paramètres dans les données en terme d'amplitude et de rayonnement en fonction de l'ouverture angulaire, et la variabilité des unités physiques des paramètres en terme de moyenne et d'écart-type. Outre ces deux points, on peut aussi rappeler que la non-linéarité intrinsèque d'un problème inverse sous-déterminé augmente avec le nombre de classes de paramètres à reconstruire, et que les données peuvent être plus ou moins reliées linéairement aux paramètres considérés.

Concernant le premier point, la sensibilité des données à un paramètre en fonction de l'angle d'ouverture détermine la résolution avec laquelle le paramètre peut être reconstruit et le degré de non-linéarité intrinsèque associé à sa reconstruction. Par exemple, la reconstruction des grandes longueurs d'ondes spatiales, généralement paramétrées par la vitesse de propagation des ondes, constitue un problème fortement non-linéaire. Ces grandes longueurs d'onde peuvent être reconstruites à partir des composantes grand-angle des données (ondes réfractées, ondes réfléchies super-critiques), auquel cas la non-linéarité résulte de la sensibilité de l'inversion à la précision du modèle de vitesse en raison du nombre élevé de longueurs d'ondes propagées (Sirgue, 2006; Pratt, 2008; Virieux & Operto, 2009). L'inversion de fréquences aussi faibles que possible constitue le meilleur remède pour réduire ces non-linéarités (Plessix *et al.*, 2010a). Alternativement, ces grandes longueurs d'ondes sont reconstruites par analyse de vitesse à partir de données de sismique réflexion, auquel cas la non-linéarité résulte de l'ambiguïté entre la vitesse et la profondeur des réflecteurs si le dispositif d'écoute n'est pas suffisamment long par rapport à la profondeur d'investigation. La reconstruction des courtes longueurs d'ondes du sous-sol à partir des ondes réfléchies représente un problème plus linéaire car étroitement lié à l'extraction de l'information contenue dans l'amplitude des ondes réfléchies, une fois que les grandes longueurs d'ondes du milieu ont été reconstruites lors d'une phase préliminaire de l'inversion. Ce découplage d'échelles associé à la phase de construction du modèle de vitesse et à l'imagerie de la réflectivité par migration constitue le fil conducteur du protocole d'imagerie sismique en sismique réflexion multitrace. L'inversion des formes d'ondes en considérant des arrivées sismiques enregistrées avec des incidences variables et des fréquences plus faibles tend idéalement à réunifier ces deux tâches au sein d'un processus intégré par un échantillonnage continu des fréquences spatiales du sous-sol.

La paramétrisation adoptée par l'inversion des formes d'ondes est fondamentale car elle influe sur la résolution avec laquelle les paramètres peuvent être imagés et sur les couplages poten-

tiels entre les différents paramètres. La résolution est contrôlée par la sensibilité des données au paramètre en fonction de l'angle d'incidence. En d'autres termes, la sensibilité est contrôlée par le diagramme de rayonnement du paramètre, qui lui-même contrôle l'amplitude des dérivées partielles du champ d'ondes modélisé par rapport au paramètre en fonction de l'angle de diffraction. Une bonne résolution est atteinte lorsque cette sensibilité est élevée pour une large gamme d'incidences. Si deux classes de paramètres rayonnent dans la même bande d'ouvertures angulaires, on peut s'attendre à un couplage de ces paramètres lors de leur reconstruction, différentes combinaisons de ces paramètres pouvant expliquer de la même manière les données. Il peut être alors judicieux de sélectionner une paramétrisation telle que les différentes classes de paramètres rayonnent dans des gammes d'ouvertures angulaires différentes. L'inconvénient de cette stratégie est de réduire la bande passante des modèles du sous-sol (en terme de fréquences spatiales) et de ce fait de dégrader la résolution potentielle de l'imagerie. J'aurai l'occasion dans la suite d'illustrer ce problème de compromis entre résolution et couplage multiparamètres avec les reconstructions respectives des paires (V_P, ρ) et (V_P, I_P) .

Par ailleurs, il est nécessaire que le paramètre considéré ait une empreinte de magnitude suffisamment forte dans les données pour que la minimisation des résidus génère une perturbation significative et stable du paramètre. Le calcul des dérivées de Fréchet $(\partial d / \partial m)$ permet d'analyser cette sensibilité des données aux paramètres, et nous savons que le terme associé à la diffraction simple dans le Hessien est composé de l'inter-corrélation de ces dérivées partielles. Le Hessien est une matrice structurée par blocs où les blocs diagonaux associés à chaque classe de paramètre contiennent toutes les dérivées secondes des données par rapport aux paramètres d'une même classe i' : $\partial^2 d / (\partial m_{i,j} \partial m_{i,k})$, où j et k indiquent deux positions dans le modèle. Les termes diagonaux d'un bloc à l'autre peuvent alors se révéler d'amplitudes très différentes selon la sensibilité des paramètres, rendant la matrice Hessienne mal conditionnée (Gouveia & Scales (1998, leur figure 12)). Il est dans ce cas nécessaire d'ajouter un terme d'amortissement (λ) sur la diagonale du Hessien pour éviter les divisions par des nombres proches de zéro. Le choix de ce terme d'amortissement est crucial : une valeur trop faible est à l'origine d'instabilités dans le modèle reconstruit, et un paramètre trop fort tue la contribution du Hessien des données.

La magnitude de l'empreinte des paramètres dans les données peut aussi conditionner la stratégie d'inversion. Une reconstruction simultanée des différentes classes de paramètres est conseillée si les paramètres ont une empreinte d'importance équivalente dans les données, alors que, dans le cas contraire, il peut être préférable de suivre une approche hiérarchique en considérant d'abord les paramètres dominants avant les paramètres de second ordre (Tarantola, 1986). Le risque associé à ce dernier choix est que le paramètre reconstruit lors la première étape absorbe les contributions des paramètres secondaires si un couplage existe entre ces paramètres.

L'unité physique détermine logiquement l'ordre de grandeur du paramètre. Lorsque nous considérons des termes de régularisation dans l'espace des modèles dans une fonction coût de type Tikhonov, il faut que l'importance de chaque paramètre soit rendue indépendante du choix de l'unité. C'est pour cela que sont introduites des normalisations par rapport à chaque paramètre via la matrice de pondération \mathbf{W}_m . La dynamique de l'échelle, c'est-à-dire son écart-type, joue aussi un rôle dans le poids du paramètre dans la partie régularisante de la fonctionnelle. Une solution possible réside dans l'utilisation d'un changement de variable tel que le logarithme du paramètre pour absorber les grands écarts-types (Tarantola, 1987). Nous discuterons de ces aspects dans la partie où sont reconstruits la vitesse, la densité et l'impédance.

Les blocs non-diagonaux du Hessien sont constitués par l'intercorrélation des dérivées partielles

des données (d) par rapport à deux classes de paramètres m_i et m_j (soit $\partial d/(\partial m_i \partial m_j)$). Ces blocs non diagonaux sont supposés corriger le gradient des couplages (ou sensibilités croisées) entre les paramètres. Je prendrai dans la suite en compte ces effets correcteurs du Hessien avec l'algorithme d'optimisation de quasi-Newton L-BFGS.

L'ensemble des points décrits déterminera le comportement de la fonction coût, qui se doit d'être localement quadratique pour que le problème inverse soit bien posé.

Je présente maintenant les résultats que nous voulons soumettre à publication dans un article subdivisé en deux parties. La première partie présente la reconstruction des paramètres visco-acoustiques (V_P, ρ, Q_P). La deuxième partie qui sera traitée dans le chapitre suivant concerne la reconstruction des vitesses d'ondes P et S dans l'approximation visco-élastique à partir des composantes d'hydrophone et de géophone.

**Multiparameter full waveform inversion of multicomponent OBC
data from the Valhall field.**

Part 1 : imaging compressional wavespeed, density and attenuation.

V. Prioux, R. Brossier, S. Operto, J. Virieux

For submission to Geophysics

5.2 Summary

Multiparameter FWI is a challenging and important issue for lithological characterization and reservoir monitoring. Mainly, the two principal issues when dealing with multiparameter inversion are [i] the variable imprint of the parameters within the data in terms of magnitude and of directivity, and [ii] the variable physical units of the parameters, in terms of mean and standard deviation. In this study we consider a 2-D frequency domain Discontinuous Galerkin method within the visco-acoustic approximation to invert the compressive velocity (V_P), the density (ρ) or the impedance (I_P), and the quality factor (Q_P , the inverse of the attenuation). We first evaluate on the synthetic Valhall case study the parametrizations (V_P, ρ) and (V_P, I_P). We conclude, contrary to what is sometimes suggested, that the first parametrization allows a better reconstruction of both the V_P , ρ and I_P parameters, firstly because it provides a broadband reconstruction of the V_P model and secondly because the small standard deviation of the density makes this deduced parameter very sensitive to small errors on the reconstructed V_P and I_P FWI models inferred from the second parametrization (using the relation $\rho = I_{P,FWI}/V_{P,FWI}$). We also show that, with the first parametrization, a hierarchical strategy where the velocity is inverted before the joint inversion of the velocity and density, allow to recover a more stable reconstruction of the density. Then, we test, on the same synthetic case, different inversion strategies to recover the following three parameters V_P , ρ and the quality factor Q_P . Simultaneous inversion tends to over-estimate the secondary parameters ρ and Q_P while the hierarchical inversion under-estimate them, but at the benefit of a gain in resolution as the inverted V_P in a first step of the inversion provides an initial correct kinematic fit. In the framework of the Tikhonov regularization, we promote the choice of an adapted damping inversion term for each parameter (we proceed by trial-and-error) in order to compensate the weaker sensitivity of the secondary parameters. Finally, we apply the simultaneous and hierarchical strategy to invert V_P , ρ and Q_P from 4C OBC wide-aperture data recorded in the

Valhall oil field. The low density contrasts at 0.6 *km* and 1.5 *km* in depths are interpreted as an accumulation of gas below barriers of claystones while the Q_P -factor show weak values down to 600 *m* in depth. A potentially interesting Q_P might be related to a gas-filled fracture at a depth of 1 *km*. Multiparameter results derived within the anisotropic visco-acoustic approximation compare quite nicely with well logs and time-domain seismic modeling, giving us confidence in the proposed procedure.

5.3 Introduction

Full waveform inversion (FWI) has become part of the seismic workflow to build high-resolution P-wave velocity model, which can be used as an improved background model for reverse time migration in complex subsalt and subbasalt environment (Sirgue *et al.*, 2010; Plessix & Perkins, 2010; Vigh *et al.*, 2011). FWI is an optimization problem, which seeks to minimize the misfit between the recorded and the modeled seismic data (Tarantola, 1984b, 1987; Pratt *et al.*, 1998; Virieux & Operto, 2009). The synthetic data are computed with the full two-way wave equation in an attempt to exploit the full information content of the data associated with diving waves, pre- and post-critical reflections, diffractions and so on. Several simplifications in the physical description of the subsurface are generally considered for seismic modeling and inversion. The most basic one considers that the subsurface is an acoustic medium parametrized by the P wavespeed only (e.g. Ravaut *et al.*, 2004; Brenders & Pratt, 2007b). These simplifications are motivated by the need to reduce the computational burden of the seismic modeling and the nonlinearity of the FWI, this nonlinearity increasing when multiple classes of parameters are reconstructed. Although high-resolution P-wave velocity models can provide useful inferences for geological interpretation, a more realistic description of the physical properties of the subsurface, in particular of the elastic properties is required for realistic reservoir and fluid characterization (Shi *et al.*, 2007).

This study is the first part of a series of two publications, which present an application of 2D multiparameter FWI to multicomponent ocean-bottom cable (OBC) data from the shallow-water Valhall field to build a visco-elastic model of the subsurface along two cables of the 3D acquisition. As above-mentioned, FWI becomes highly nonlinear when multiple classes of parameters with contrasted imprint in the data are reconstructed. Hierarchical inversions of the data components and hierarchical update of the multiple classes of parameter may be viewed to reduce this non linearity (Tarantola, 1986; Sears *et al.*, 2008; Jeong *et al.*, 2012). In this study, we will discuss the feasibility of the reconstruction of the P wavespeed, the density or impedance and the Q_P -factor from the hydrophone component in the visco-acoustic approximation. A key issue here is to choose a suitable parametrization of the subsurface during FWI, which allows one to manage the variable sensitivity of the data to the parameters as well as the trade-off between parameters. In the second paper, we shall use the subsurface models built in this study as initial or background models to jointly update the P and S wavespeeds from the hydrophone and the geophone components.

Only few studies have been dedicated to the reconstruction of density and Q_P -factor by FWI, although it is acknowledged that these two parameters can have a significant footprint on wave amplitudes (Jeong *et al.*, 2012; Hicks & Pratt, 2001; Malinowski *et al.*, 2011). Generally speaking, it is acknowledged that seismic reflection imaging can be subdivided into two distinct tasks, which honor the scale separation between the velocity macro-model and the reflectivity :

reconstruction of the large wavelengths of the P-wave velocity followed by the imaging of the reflectivity parametrized by the impedance (Tarantola, 1986; Métivier *et al.*, 2012b). While the velocity-impedance parametrization is natural in seismic reflection processing, one might wonder whether this parametrization is the most suitable one for FWI of wide-aperture data, during which the scale separation is not used as a continuous spectrum of wavenumbers can be imaged (Pratt & Worthington, 1990; Neves & Singh, 1996; Sirgue & Pratt, 2004). The first topic tackled in this study is to assess whether the velocity-impedance parametrization is more suitable than the velocity-density one for FWI of wide-aperture data. As we shall see, two criteria to select the best parametrization are related to the trade-off between the two parameter classes and the resolution with which the two parameter classes can be reconstructed.

Only few case studies of Q_P -factor imaging by FWI have been presented in marine and land environments (Hicks & Pratt, 2001; Malinowski *et al.*, 2011). Attenuation has generally a much smaller imprint in the data than wavespeed, in particular, in marine environment. Therefore, the second topic addressed by this study is related to the design of a robust hierarchical optimization algorithm, which is able to handle multiple classes of parameter with variable footprint in the data.

In this study, FWI is performed in the frequency domain using a quasi-Newton optimization and a Tikhonov regularization. In the first part of the study, we review the key methodological aspects, which are useful for multiparameter FWI. Two of them concern the optimization algorithm, which should take into account the effects of the Hessian during inversion, and the regularization, which should help to reduce the nonlinearity of the inversion without affecting the ability of the FWI to update parameters with a smaller imprint in the data. Second, we discuss with a realistic synthetic example representative of Valhall which of the velocity-impedance and velocity-density parametrizations is the most suitable one for FWI of wide-aperture data. Finally, we propose a hierarchical approach for imaging the compressional wavespeed, density and Q_P -factor from the synthetic Valhall model and the real OBC data from Valhall.

5.4 Visco-acoustic full waveform inversion

In this study, we seek to reconstruct the vertical P wave speed, the density or impedance and the Q_P -factor (the inverse of the attenuation) by visco-acoustic FWI from hydrophone data. The subsurface is assumed to be a visco-acoustic vertically transverse isotropic (VTI) medium. Seismic modeling is performed in unstructured triangular meshes with a frequency-domain velocity-stress finite-element discontinuous Galerkin method (Brossier *et al.*, 2010b).

Full waveform inversion (FWI) is performed in the frequency domain by local optimization, where the gradient of the misfit function is computed with the adjoint-state method (Plessix, 2006; Chavent, 2009). A forward-modeling operator based on the second-order wave equation for particle velocities and piecewise constant interpolation function (P0) in the finite-element method is used to build the diffraction kernel of the FWI from self-adjoint operators, while the incident and the adjoint wavefields are accurately computed with a first-order velocity-stress formulation of the wave equation with either P0, P1 (linear) or P2 (quadratic) interpolation functions (Brossier, 2011a). Of note, the velocity-stress wave equation allows us to introduce in a natural way the pressure during seismic modeling for both the computation of the incident and adjoint wavefields.

The misfit function $\mathcal{C}(\mathbf{m})$ is given by

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^\dagger \mathbf{W}_d \Delta \mathbf{d} + \frac{1}{2} \sum_{i=1}^{N_p} \lambda_i (\mathbf{m}_i - \mathbf{m}_{prior_i})^\dagger \mathbf{W}_{m_i} (\mathbf{m}_i - \mathbf{m}_{prior_i}), \quad (5.2)$$

where $\Delta \mathbf{d}$ denote the hydrophone data-residual vector, which is weighted by the operator \mathbf{W}_d . The matrix \mathbf{W}_d can weight data residuals according to the standard error and/or according to source-receiver offset (Ravaut *et al.*, 2004; Operto *et al.*, 2006). The multiparameter subsurface model is denoted by $\mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_{N_p})$, where N_p denotes the number of parameter classes. In this study, three classes of parameter are considered : the vertical compressional wave speed (V_{P0}), the density (ρ) or the impedance (I_P) and the quality factor (Q_P). A Tikhonov regularization is applied to each class of parameter i through a roughness operator \mathbf{W}_{m_i} , which forces the difference between the model \mathbf{m}_i and a prior model \mathbf{m}_{prior_i} to be smooth. The scalars λ_i control the weight of the data-space misfit function $\frac{1}{2} \Delta \mathbf{d}^\dagger \mathbf{W}_d \Delta \mathbf{d}$ relatively to the model-space misfit functions $\frac{1}{2} (\mathbf{m}_i - \mathbf{m}_{prior_i})^\dagger \mathbf{W}_{m_i} (\mathbf{m}_i - \mathbf{m}_{prior_i})$. Of note, the scalar λ_i can be adapted to each parameter class i , that is helpful to account for the variable sensitivity of the data to each parameter class, as we shall see.

The solution for the perturbation model, that minimizes the misfit function at iteration k is given by

$$\Delta \mathbf{m}^{(k)} = \Re \left(\widehat{\mathbf{W}}_m^{-1} \mathbf{J}^{(k)t} \mathbf{W}_d \mathbf{J}^{(k)*} + \widehat{\mathbf{W}}_m^{-1} \left(\frac{\partial \mathbf{J}^{(k)t}}{\partial \mathbf{m}^t} \right) (\Delta \mathbf{d}^{(k)*} \dots \Delta \mathbf{d}^{(k)*}) + \Lambda \right)^{-1} \Re \left(\widehat{\mathbf{W}}_m^{-1} \mathbf{J}^{(k)t} \mathbf{W}_d \Delta \mathbf{d}^{(k)*} + \Lambda (\mathbf{m}^{(k)} - \mathbf{m}_{prior}) \right), \quad (5.3)$$

where Λ is a block diagonal damping matrix :

$$\Lambda = \begin{pmatrix} \lambda_1 \mathbf{I}_M & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & \lambda_{N_p} \mathbf{I}_M \end{pmatrix}, \quad (5.4)$$

and \mathbf{I}_M is the identity matrix of dimension M , where M denotes the number of nodes in the computational mesh. In equation 5.3, \Re is the real part of a complex number. The matrix $\widehat{\mathbf{W}}_m$ is a $N_p \times N_p$ block diagonal matrix, where each block is formed by the \mathbf{W}_{m_i} matrices. The term, whose inverse must be computed in equation 5.3 is the full Hessian, which contains 3 terms. The first one aims to correct the gradient for linear effects such as the limited bandwidth of the source, the limited spread of the acquisition geometry, and the geometrical spreading of the data. The second term corrects the Hessian for double-scattering effects in nonlinear inverse problems (Pratt *et al.*, 1998). The third regularization term damps the deconvolution action of the first two terms of the Hessian to improve the conditioning of the Hessian matrix. The term on the right-hand side of the inverse of the Hessian in equation 5.3 is the gradient of the misfit function, which is composed of two terms : the first one represents the contribution of the data and is formed by the zero-lag correlations of the partial derivative wavefields at the receiver positions with the data residuals, while the second term is the model-space regularization term.

It is worth reminding that the partial derivative wavefields satisfy

$$\mathbf{B}(\omega, \mathbf{m}(\mathbf{x})) \frac{\partial \mathbf{v}}{\partial m_i} = - \frac{\partial \mathbf{B}(\omega, \mathbf{m}(\mathbf{x}))}{\partial m_i} \mathbf{v}, \quad (5.5)$$

where $\mathbf{B}(\omega, \mathbf{m}(\mathbf{x}))$ is the forward-problem operator (in our case, the second-order wave equation operator for particle velocities) and \mathbf{v} is the incident particle velocity wavefield. The right-hand side of equation 6.4 is the secondary virtual source of the partial derivative wavefield, the spatial support and the temporal support of which are centered on the position of the diffractor m_i and on the arrival time of the incident wavefield at the diffractor m_i , respectively (Pratt *et al.*, 1998). The diffraction or radiation pattern of this virtual source is given by $\frac{\partial \mathbf{B}(\omega, \mathbf{m}(\mathbf{x}))}{\partial m_i}$ and gives some insights on the sensitivity of the data to the parameter m_i as a function of the diffraction (or aperture) angle.

We use the quasi-Newton L-BFGS optimization algorithm to solve equation 5.3 (Nocedal, 1980; Nocedal & Wright, 1999). The L-BFGS algorithm computes recursively an approximation of the product of the inverse of the Hessian with the gradient, from few gradients and few solution vectors from the previous iterations. As preconditioner of this iterative search, we use a diagonal approximation of the approximate Hessian (the linear term) damped by the Λ matrix,

$$PreH = \left(\widehat{\mathbf{W}}_m^{-1} \text{diag} \left\{ \mathbf{J}^{(k)t} \mathbf{W}_d \mathbf{J}^{(k)*} \right\} + \Lambda \right)^{-1}. \quad (5.6)$$

Of note, the smoothing operator $\widehat{\mathbf{W}}_m^{-1}$ is applied to the first two terms of the Hessian in equation 5.3 for implementation convenience, although another implementation consists of using $\widehat{\mathbf{W}}_m$ in the third term of the Hessian. Similar comment applies to the gradient term, where the operator $\widehat{\mathbf{W}}_m$ or its inverse can be applied to the first or second terms of the gradient accordingly. Our implementation allows the preconditioner of the L-BFGS algorithm, equation 6.5, to be diagonal and hence easy to invert because the matrix Λ is diagonal unlike $\widehat{\mathbf{W}}_m$.

The smoothing operators $\mathbf{W}_{m_i}^{-1}$ are exponential function given by :

$$\mathbf{W}_{m_i}^{-1}(z, x, z', x') = \sigma_i^2(z, x) \exp\left(\frac{-|x - x'|}{\tau_x}\right) \exp\left(\frac{-|z - z'|}{\tau_z}\right), \quad (5.7)$$

where τ_x and τ_z denote the horizontal and vertical correlation lengths, defined as a fraction of the local wavelength. The coefficient σ_i represents the standard error and is scaled to the physical unit of the parameter class i . An exponential function is used for $\mathbf{W}_{m_i}^{-1}$ because its inverse in the expression of the misfit function can be computed analytically (Tarantola (1987), section 7.21, pages 308-309).

The frequency-domain FWI is performed following a conventional multiscale approach, which proceeds from the inversion to the low frequencies to the higher ones (Pratt & Worthington, 1990; Sirgue & Pratt, 2004). A second level of data preconditioning with respect to time can be implemented through time damping of seismograms with a decaying exponential function (Brossier *et al.*, 2009b; Shin & Cha, 2009). This time damping is implemented in frequency-domain seismic modeling with complex-valued frequencies, the imaginary part of which controls the amplitude decay with time. This is shown by the expression of the Fourier transform of the damped function $p(t)e^{-\frac{(t-t_0)}{\tau}}$:

$$p(\omega + i/\tau)e^{\frac{t_0}{\tau}} = \int_{-\infty}^{+\infty} p(t)e^{i\omega t} dt, \quad (5.8)$$

where τ will be referred to as the time-damping factor in the following.

The aim of this data preconditioning is to progressively inject more complex or more resolving late-arriving phases in the inversion to reduce the risk of cycle skipping. An application of this heuristic approach to land data for the joint update of the compressional and shear wavespeeds by joint inversion of body waves and surface waves is presented in Brossier *et al.* (2009b).

5.5 Density versus impedance imaging

5.5.1 FWI parametrization and radiation patterns

It is possible to parametrize the subsurface with different combination of parameters for seismic modeling and FWI (Forgues & Lambaré, 1997; Plessix & Cao, 2011). Of note, the parametrizations used for seismic modeling and for waveform inversion can be different provided that a transformation exists to move from one parametrization to the other one. In the framework of non linear inversion, parameters of a given parametrization can be related non linearly to the parameters of an other parametrization. As mentioned by Tarantola (1986), the choice of the parametrization is not neutral in the sense that it will govern how the amplitude of a wave scattered by a diffractor associated with one model-parameter perturbation varies with the diffraction (or aperture) angle (Tarantola, 1986). This scattered wave corresponds to the partial derivative of the wavefield with respect to the parameter located at the diffractor point, equation 6.4. In the following, the radiation pattern of the secondary source of the partial derivative wavefield at the diffractor point, right-hand side in equation 6.4, will be referred to as the radiation pattern of the parameter with a slight abuse of language. This radiation pattern gives some insights on how the sensitivity of the data to the parameter changes as a function of the aperture angle. Since the aperture angle is closely related to the wavenumbers injected in the subsurface model during waveform inversion (e.g. Sirgue & Pratt, 2004), the radiation pattern gives in fact some insights on the resolution with which a given parameter class can be reconstructed. The strength of this secondary source gives some additional insight on the sensitivity of the data to the parameter relatively to the other parameters involved in the parametrization. Parametrization analysis of FWI aims at finding the parametrization, which leads to the best sensitivity of the data to each class of parameters with respect to the aperture angles for high-resolution imaging, while keeping under control the potential trade-off between parameters.

Analytical expression of the visco-acoustic/elastic radiation pattern were developed asymptotically in the framework of the ray+Born approximation in Wu & Aki (1985); Tarantola (1986); Ribodetti & Virieux (1996); Forgues & Lambaré (1997). Alternatively, they can be estimated numerically by computing the partial derivative wavefields in a finite difference sense, as shown in Malinowski *et al.* (2011) and Gholami *et al.* (2011b). In this study, we shall follow this numerical finite-difference approach and derive the radiation pattern for the three different parametrizations : (V_P, ρ) , (V_P, I_P) and (I_P, ρ) (Figure 5.3). To do so, we computed an incident monochromatic wavefield in a homogeneous model, that is subtracted to the wavefield computed in the same model in which we added a point-diffractor associated with one parameter perturbation in the center of the mesh. The resulting wavefield represents the wavefield scattered by the parameter perturbation in all directions for a given parametrization or, in other words, the partial derivative of the wavefield with respect to the model parameter located at the point diffractor. During this simulation, the other parameters of the parametrization are

INVERSION VISCO-ACOUSTIQUE DES FORMES D'ONDES - RECONSTRUCTION DE
LA VITESSE P, DE LA DENSITÉ OU L'IMPÉDANCE, ET DE L'ATTÉNUATION

Figure 5.3	Parametrization	Inclusion of	Value in the	V_P (m/s)	ρ (kg.m ⁻³)	I_P (kg.m ⁻² .s ⁻¹)
(a-f)	All	-	background	4000	2000	8.10^6
(a)	(V_P, ρ)	V_P	inclusion	8400	2000	$16.8.10^6$
(b)		ρ		4000	4200	$16.8.10^6$
(c)	(V_P, I_P)	V_P		8400	952	8.10^6
(d)		I_P		4000	4200	$16.8.10^6$
(e)	(I_P, ρ)	I_P		8400	2000	$16.8.10^6$
(f)		ρ		1905	4200	8.10^6

TABLE 5.1 – Values of the physical parameters in the homogeneous medium and the inclusion to obtain the diffracted field by an inclusion for different parametrizations.

left unchanged, and this procedure is repeated for each parameter class of the parametrization. The amplitude variations of the partial derivative wavefield around the point diffractor give some insight on the radiation pattern of the model parameter as a function of the aperture angle. Indeed, the radiation pattern of a given parameter class (for example, V_P) can change as a function of the parametrization, as we show below.

Table 5.1 outlines the values of the model parameters, that are used in the computation of the radiation patterns. The source is a Dirac function and the frequency is 10 Hz. A first conclusion is that the radiation pattern of the P-wave velocity parameter is isotropic in the (V_P, ρ) parametrization, while scattering of the P-wave velocity diffractor is only significant at large aperture angles in the (V_P, I_P) parametrization. A second conclusion is that the radiation pattern of the P-wave velocity and density overlap at small and intermediate apertures. This suggests that many combination of P-wave velocity and density can allow one to match equally well short-spread reflections, hence highlighting the trade-off between these two parameters. In contrast, the P-wave velocity and the impedance scatter waves at large and short aperture angles, respectively, when the (V_P, I_P) parametrization is considered. This highlights that the imaging of the long wavelengths are closely related to the reconstruction of the P-wave velocities from the diving waves and super-critical reflections, while the short wavelengths of the subsurface are closely related to impedance, which governs the short-spread reflections from discontinuities. This justifies the conventional workflow of seismic imaging based on the scale separation underlying the velocity model building task and migration. Since the radiation patterns of the P-wave velocity and impedance do not overlap significantly, no trade-off between these parameters is expected during FWI.

Of note, the reconstruction of the P-wave velocity from the (V_P, I_P) parametrization is expected to be smoother than the one obtained with the (V_P, ρ) parametrization. This is a criterion that must be considered to select the best subsurface parametrization. The maximum amplitudes of the 3 parameters with selected units, whatever the parametrization is, are of the same order of magnitude, which suggests that the sensitivity of the data to these parameters is significant (Figure 5.3). Therefore, the main issue here is to manage the trade-off between parameters rather than to manage parameters with contrasted imprint in the data.

In the following section, we discuss with a realistic synthetic case study which parametrization provides the most reliable reconstruction.

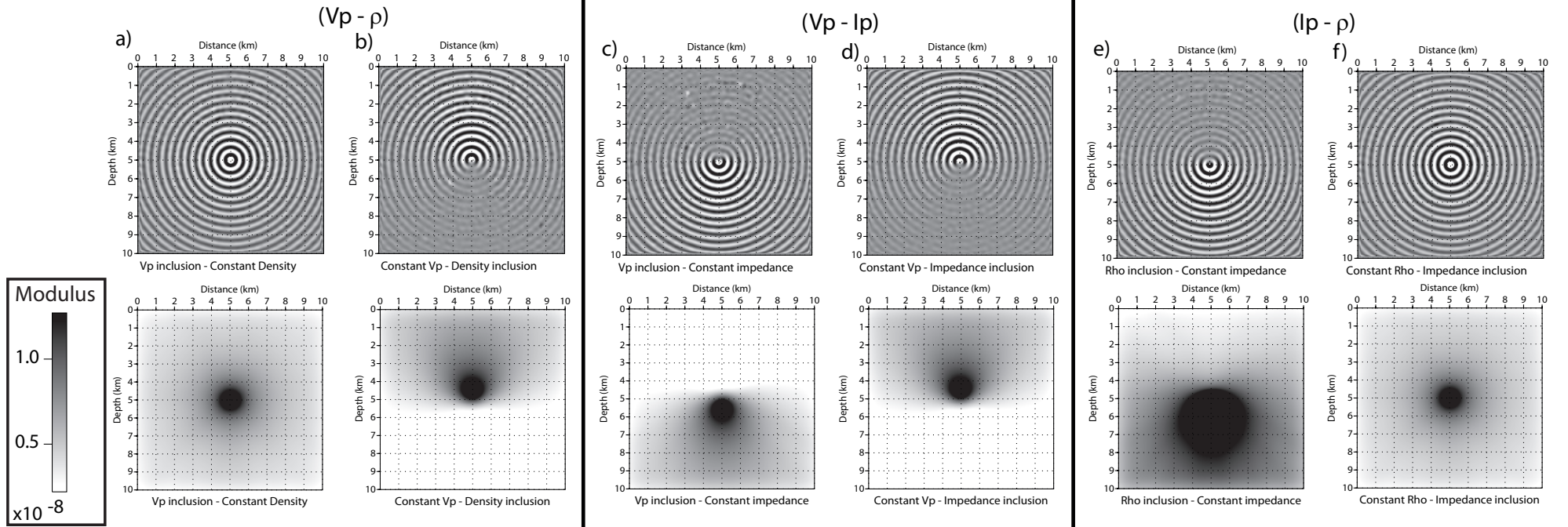


FIGURE 5.3 – Wavefields scattered by a point diffractor located in the center of the medium for the three parametrizations : (V_P, ρ) , (V_P, I_P) and (I_P, ρ) . Amplitude variations of these wavefields around the diffractor give some insights on the radiation pattern of each model parameter for a given parametrization. Top and bottom figures are respectively the real part and the modulus of the scattered wavefields.

5.5.2 Synthetic Valhall case study

5.5.2.1 Models, acquisition and full waveform inversion setup

We assess with the realistic synthetic Valhall case study the feasibility of acoustic full waveform inversion (FWI) for the two parametrizations (V_P, ρ) , and (V_P, I_P) . Implementation of the (V_P, ρ) parametrization was first implemented in the code of Brossier (2011b). Alternative parametrization can be easily implemented from an existing one using the chain rule during the computation of the radiation patterns.

We consider in all of the following tests a shallow-water medium (70 m water depth), that is 16 km long and 5.2 km deep. We consider an OBC-like sea bottom acquisition with a maximum offset of 16 km and a source and receiver spacing of 50 m. Depth of the explosives sources and of the hydrophones is 6 m and 71 m, respectively. Seismic modeling is performed on a hybrid P0-P1 triangular mesh where the P0 interpolation is set below 160 m (see Brossier (2011b); Prioux *et al.* (2011)). Modeling is performed under the isotropic approximation. Five frequencies [3.5, 4, 5, 6.1, 7.1] Hz are successively inverted during FWI. The source excitation is estimated for each shot gather at each iteration of the FWI with a L2 norm following the formula of Pratt (1999). The data functional for FWI is based on the L1 norm. Maximum number of iterations per frequency is 30 (10 iterations per damping are performed when 3 damping are used per frequency). The damping coefficients λ_i in the misfit function, equation 6.1, are the same for the P-wave velocity, the density and the impedance, and are equal to $4 \cdot 10^{-18}$. This is justified because all the parameter classes considered in this study have a footprint of similar magnitude in the data (Figure 5.3).

The true Valhall models for P-wave velocity, density and impedance, which were provided by the BP company, are shown in Figure 5.4(a-d). Initial models for FWI are built by Gaussian smoothing of the true models below the sea level, with a horizontal and vertical correlation lengths of 500 m, as shown in Figure 5.4(e-h). The Q_P -factor model was built by picking on the velocity model some geological interfaces. The values of Q_P are set to 80 where the velocity is smaller than 1600 m/s above 730 m in depth and smaller than 1800 m/s below 730 m in depth, and are related to the gas layers. Small Q_P values (of 60) above 730 m are associated with soft quaternary sediments.

To perform a quality control of the FWI models, we compute a relative percentage of error for the models ($\xi(\mathbf{m})$) and for the seismograms ($\xi(\mathbf{s})$), defined by

$$\xi(\mathbf{m}_{cal}) = \frac{100}{N_x \cdot N_z} \sum_{i=1}^{N_x \cdot N_z} \frac{|m_{cal,i} - m_{true,i}|}{m_{true,i}}, \text{ and } \xi(\mathbf{s}_{cal}) = \frac{100}{N_t \cdot N_s} \sum_{i=1}^{N_s \cdot N_t} \frac{|s_{cal,i} - s_{true,i}|}{s_{true,i}}, \quad (5.9)$$

where s_{cal} denotes a time-domain shot gather. The number of time samples and the number of seismograms per shot gather are denoted by N_t and N_s , respectively, while N_x and N_z denote the number of nodes in the subsurface model. We compute time-domain seismograms from the frequency-domain modeling for 3 shot gathers located at 2 km, 8 km and 14 km in distance. Of note, shot gathers 1 and 3 have a maximum offset of 14 km, while the shot gather 2 mainly records short offset waves with a maximum offset of 8 km.

To fairly compare the mean error in the models and to be independent of the range of variation of each parameter, we also define the normalized mean error $\xi_{mean,n}(\mathbf{m})$, using for

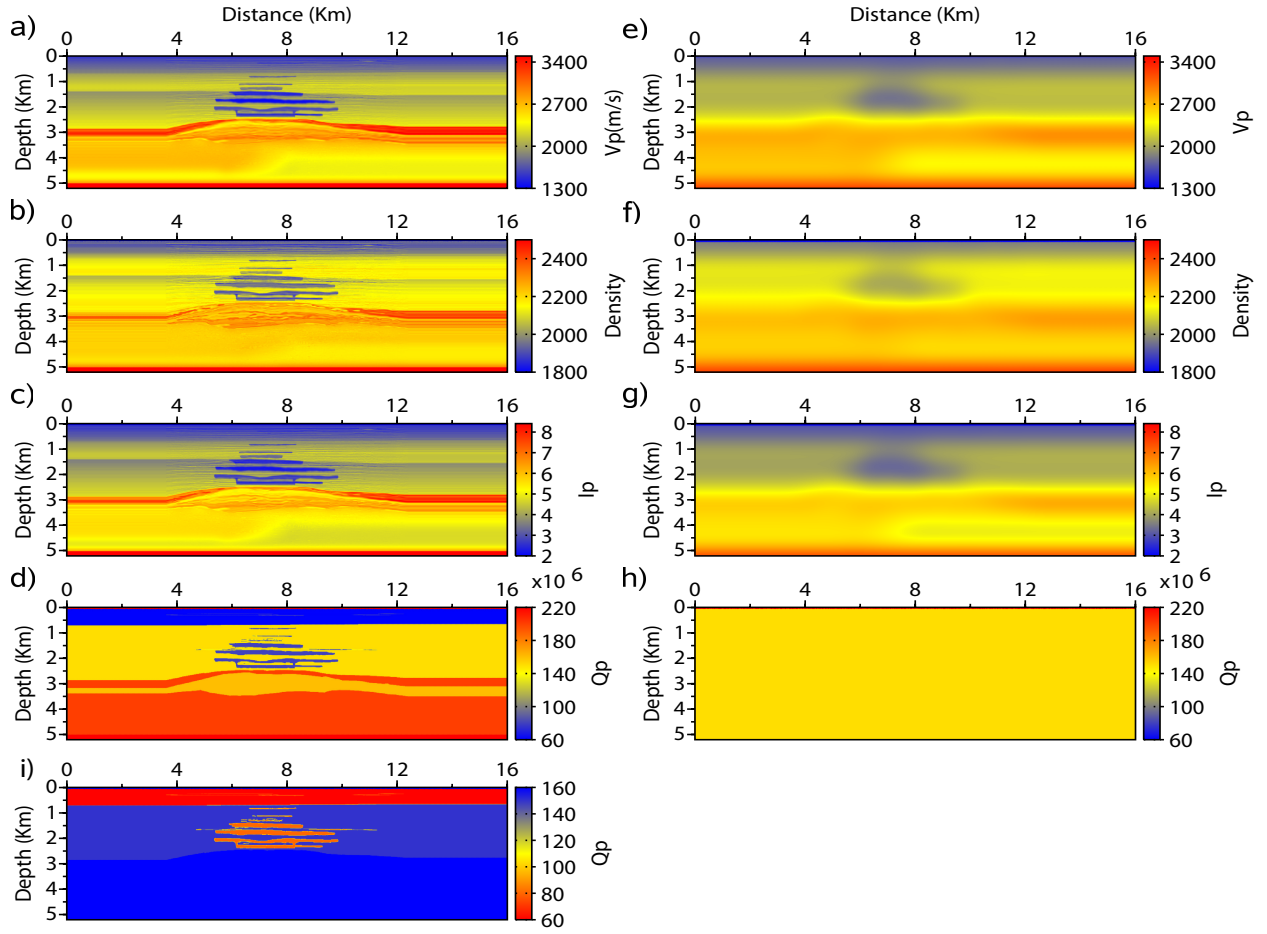


FIGURE 5.4 – True (a-d) and corresponding starting models (e-h) for the velocity (e), the density (f) and the impedance (g). The Q_P -factor background model is homogeneous with $Q_P = 150$ and 1000 in the subsurface and in the water, respectively (h). (i) True Q_P -factor model.

each parameter a normalization factor set to the higher error percentage among all the tests (in blue color in the Table 5.3).

The model and data errors for all of the tests presented in the following are outlined in Table 5.2

5.5.2.2 FWI results : joint reconstruction of (V_P, ρ) and (V_P, I_P)

We first focus on the reconstruction of the parameter classes (V_P, ρ) versus (V_P, I_P) . Therefore, no attenuation is considered in this study (Q_P is homogeneous in both the true and background models and is equal to 1000). The results of the joint reconstruction of the velocity and the density with the (V_P, ρ) parametrization are shown in Figures 5.5(a-c) and 5.6(a-c)), while the results of the joint reconstruction of the velocity and the impedance with the (V_P, I_P) parametrization are shown in Figures 5.5(d-f) and 5.6(d-f). Three successive time dampings, equation 5.8, are used during each of the mono-frequency inversions ($\tau = 1, 3, 10s$).

For both parametrizations, the data error for the shot gather located at 8 km in distance is

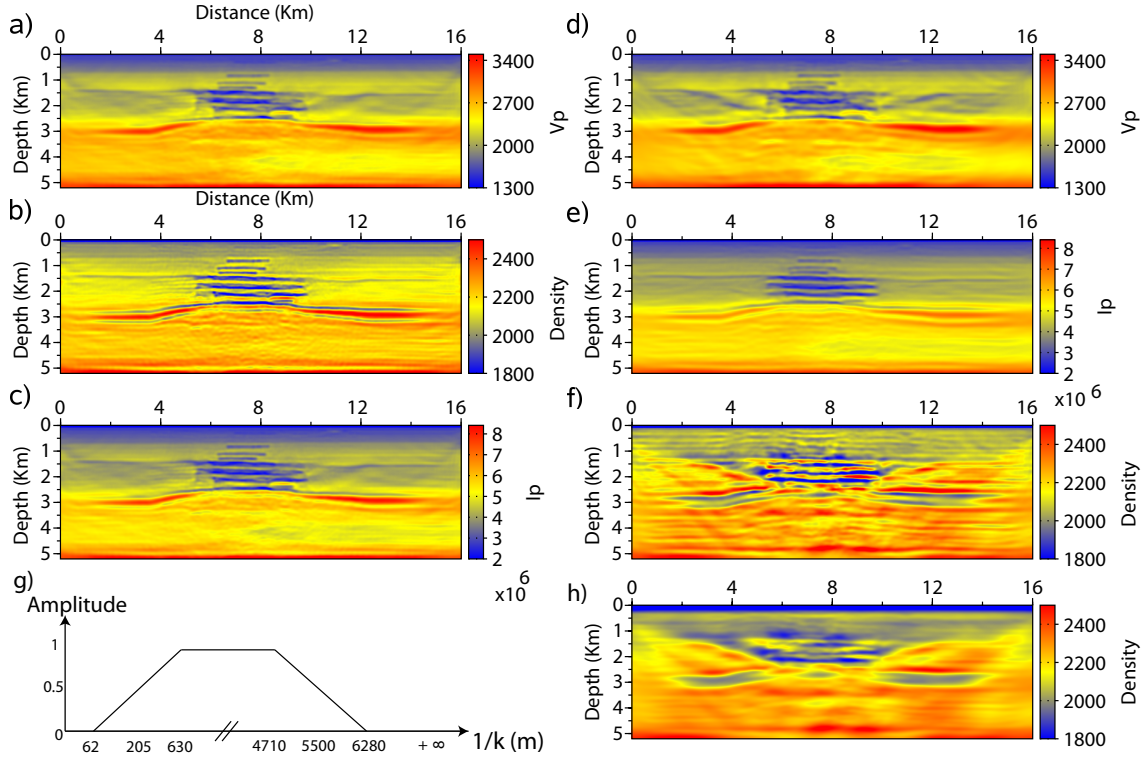


FIGURE 5.5 – Joint update of two parameters. (a-b) Final FWI V_P and ρ models for the parametrization of type 1, and (c) I_P model inferred from those two models ($I_P = \rho.V_P$). (d-e) Final FWI V_P and I_P models for the parametrization of type 2, and (f) ρ model inferred from those models ($\rho = I_P/V_P$). (g) $k_x - k_z$ domain filtering law applied to the V_P and the I_P models shown in (d) and (e), that were used to infer the ρ model shown in (h).

higher because waves at short offsets have higher amplitudes than at long offsets, and because FWI tends to be dominantly controlled by the wide apertures in the data, the misfit of which is reduced first (Table 5.2). The model error highlights not only the model inaccuracies associated with convergence towards a local minimum but also the resolution with which a model is reconstructed according to the radiation pattern of the parameter. Both sources of errors should be taken into account in the appraisal of the models. For example, the percentage of error for the density starting model (1.94 %) is smaller than the error for the FWI density model (2.54%), despite the numerous structures that FWI created. In the other hand, the percentage of error on the seismograms computed in the starting models is almost 3 times higher than the one inferred from the FWI models, hence showing the relevance of this last criterion when we want to compare models characterized by a different resolution.

The P-wave velocity reconstruction for the (V_P, ρ) parametrization is better resolved than the one inferred from the (V_P, I_P) parametrization, in which some artifacts are shown down to 2.5 km in depth (Figure 5.5). These inaccuracies might be related to the narrow radiation pattern of the P-wave velocity when the (V_P, I_P) parametrization is used, that prevents a broadband reconstruction of the P-wave velocity model. The lack of high wavenumbers in the P-wave velocity model inferred from the (V_P, I_P) parametrization is more obvious in the upper

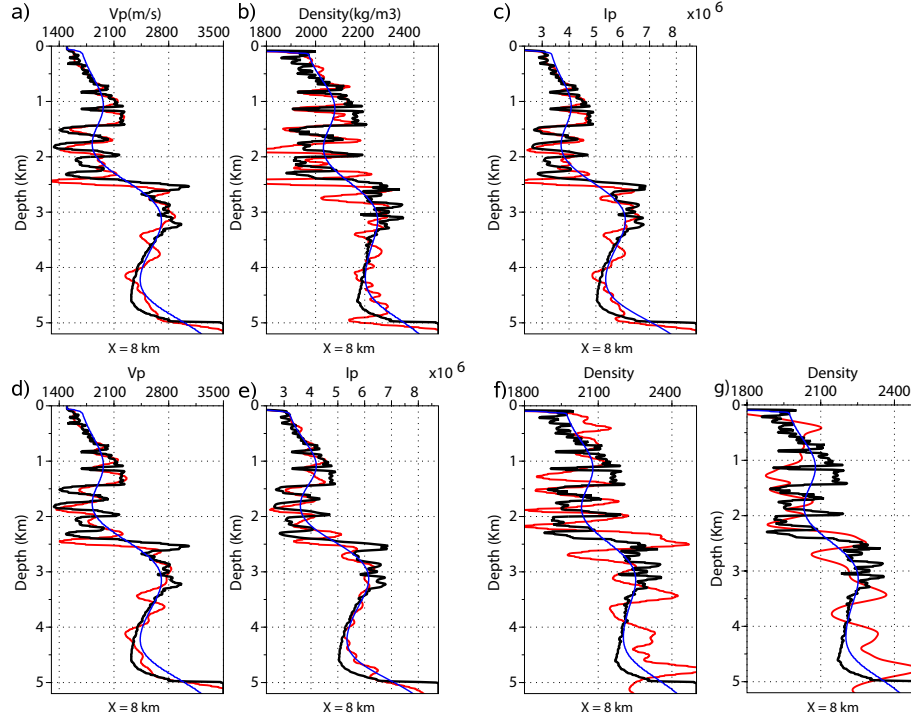


FIGURE 5.6 – (a-b) Logs extracted from the final FWI V_P and ρ models for the (V_P, ρ) parametrization. (c) Impedance log inferred from the (V_P, ρ) models shown in (a). (d-e) Logs extracted from the final FWI V_P and I_P models for the (V_P, I_P) parametrization. (f) Density log inferred from the two previous models. The black curve denotes the true model, the blue line the starting models obtained by smoothing the true models with a correlation length of 500 m, and the red line the FWI model. (g) Filtered version of the (f) profile, extracted from the model shown in Figure 5.5h.

structure, because the aperture illumination is broader in the shallow part making the filtering applied by the directivity of the virtual sources more visible.

The density model shows a sharp reconstruction of the main reflector (Figure 5.5b) with locally overestimated or underestimated perturbations (Figure 5.6b). However, the impedance model reconstructed from the FWI (V_P, ρ) models shows a good agreement with the true impedance model, except at the reservoir level where a mispositioning of the reflector at 2.5 km in depth is shown. The good agreement between the true and the imaged impedance models suggest that the errors in the amplitudes of the density perturbations might result from the trade-off between velocity and density at short apertures, these trade-off effects being canceled when these two parameters are recombined to build the impedance.

The impedance model inferred from the (V_P, I_P) parametrization shows a good agreement with the true impedance model (Figures 5.5e and 5.6e). However, the density model inferred from the (V_P, I_P) models does not show a good agreement with the true density model.

The impedance model inferred from the velocity and the density FWI models obtained with the (V_P, ρ) parametrization (Figure 5.5(c)) is of better quality than the impedance model found by FWI with the (V_P, I_P) parametrization (Figure 5.6(e)), with an impedance error 0.32 % higher in the last case (Table 5.2). Furthermore, the density model inferred from the impedance and velocity FWI models obtained with the (V_P, I_P) parametrization is very unstable (Figure 5.5(f)), with a model error being almost 2 % higher than the density model inferred from the (V_P, ρ) parametrization. One may argue that those instabilities are related to the fact we divide two models of different resolution, the velocity model inferred from the (V_P, I_P) parametrization suffering from a deficit of high wavenumbers. To test this hypothesis, we applied a $k_x - k_z$ filtering (whose shape is defined Figure 5.5g) to the velocity and impedance FWI models in order to force their resolution to be similar. We used those two filtered models to compute the corresponding density models shown with their profiles in Figures 5.5h and 5.6g. Although smoother, the resulting density model remains inaccurate, in particular in the first 1.5 km in depth and below the reservoir level. Error computed in this density model is 4.59 % against 4.51 % for its original version. A more plausible reason to explain these instabilities can be related to the relative range of variations of each parameter, defined by : $(value\ max - value\ min)/value\ max$. This quantity is equal to 60 %, 26 % and 73 % for the velocity, density and impedance synthetic models respectively. The order of magnitude of these percentages shows the same trend than the magnitudes of errors computed in each of the starting models (Table 5.2). This percentage is higher for the impedance because this is the product of the two other parameters, which vary within the same direction at all depths, as shown in the true logs of velocity and density (black lines in Figure 5.6(a-b)). This implies that a small error in the impedance or the velocity reconstruction will translate into a much higher relative error for the density inferred from those two models. In contrast, impedance inferred from the velocity and the density is not very sensitive to an error in the density. This statement will apply for all of the subsurface models where velocity and density vary in the same direction, that is the most common scenario in geological targets.

We illustrate with this example that the inference of one parameter from a nonlinear combination of other parameters can be an unstable process, when the inferred parameter has a small range of variations compared to the parameters involved in the inversion. The fact we fail to derive a stable density model with parametrization 2 is without consequence if one is more interested in the impedance than in the density.

5.5.2.3 FWI results : hierarchical reconstruction of two parameters

In the following, we assess whether a hierarchical reconstruction of parameters is more robust than a simultaneous update of the parameters. In our hierarchical approach, we update the P-wave velocity with a single time damping τ of 1 s during a first inversion to cancel most of the residuals at large aperture angles, and in a second step we jointly update the P-wave velocity and the density or the impedance with three time dampings to reduce the remaining residuals at short aperture angles, for which density and impedance have a significant imprint. During the second step, we jointly update the density or impedance and the velocity, as there can be a trade-off between those parameters. During the first inversion, the P-wave velocity is updated with the (V_P, ρ) parametrization because the radiation pattern of the P-wave velocity is broader when this parametrization is used. The trade-off between the P-wave velocity and density at short apertures should not impact the imaging of the wave speed because the inversion is limited

Test No.	Test		$\xi(\mathbf{m})$			$\xi_{mean,n}(\mathbf{m})$	$\xi(s_1)$. 10^{-5}	$\xi(s_2)$. 10^{-5}	$\xi(s_3)$. 10^{-5}	$\xi_{mean}(\mathbf{s})$. 10^{-5}
	STEP 1	STEP 2	V_P	ρ	I_P					
-	$V_{P,0}, \rho_0, I_{P,0}$		5.91	1.94	7.73	0.98	5.04	6.80	5.36	5.73
1	(V_P, ρ)		4.31	2.54	6.23	0.82	1.61	2.80	1.82	2.08
2	(V_P, I_P)		4.49	4.51	6.55	1	1.80	3.26	2.1	2.39
3	V_P		4.59				1.76	2.95	1.92	2.21
4	V_P	(V_P, ρ)	4.28	2.17	5.85	0.77	1.50	2.65	1.71	1.95
5	V_P	(V_P, I_P)	4.38	2.77	5.98	0.83	1.53	2.61	1.71	1.95

TABLE 5.2 – Table giving errors for the models ($\xi(\mathbf{m})$) and for the seismograms ($\xi(\mathbf{s})$) in the starting models and in the FWI models obtained with the simultaneous or the hierarchical strategies, to recover the velocity, the density and the impedance. The smallest and the biggest errors by column are denoted by the red and blue numbers respectively.

to the wide apertures through the use of aggressive time damping. The P-wave velocity model inferred from this first inversion step is shown in Figures 5.7a and 5.8(a,d) (by the blue line). The resulting model is of satisfying quality, although its resolution is poorer than that of the model obtained during the joint reconstruction of velocity and density using three dampings (Figures 5.5(a) and 5.6a). The final FWI P-wave velocity model of the first inversion step is used as the initial model for the second inversion step. The initial density model is the smoothed version of the true model when the (V_P, ρ) parametrization is used during the second inversion step. The product of velocity inferred from the first inversion step with the smoothed density model provides the starting impedance model for the inversion performed with the (V_P, I_P) parametrization (Figure 5.7e (blue line)). Final models and logs of the second inversion step for the two parametrizations can be compared in Figures 5.7 and 5.8. The FWI density model obtained with the (V_P, ρ) parametrization (Figure 5.7b) shows weaker perturbations than the one obtained during the joint update of the P-wave velocity and density (Figure 5.5b), that shows that many combinations of velocity and density models can match the data equally well. In addition, we show that the density is under-estimated down to 1.4 km in depth, and is polluted by ringing artifacts in the deep structure, which can result from the narrow aperture illumination at depth. The impedance model inferred from the P-wave velocity and density model shows a good agreement with the exact model, except between 2 and 2.5 km in depth (Figures 5.7d and 5.8c).

We show the difference between the seismograms computed in the true models and the models obtained close of the first inversion step (the first FWI V_P model and the smooth initial density model) in Figure 5.9d. The first step of the inversion minimized most of the residuals in the data, hence highlighting the risk that the velocity accounts in a first step for the effects of the density. In the other hand, the seismograms computed in the final models of the second step (with the (V_P, ρ) parametrization) and its residuals (Figure 5.9(e-f)) show that this second step allowed us to improve the fit of the seismograms in particular for the short-offset reflected waves, as confirmed by a time error 0.26 % smaller (Table 5.2).

During the second step of the inversion performed with the (V_P, I_P) parametrization, the FWI velocity and impedance models did not bring important perturbations as their respective starting models were already quite close to the true model (Figure 5.8(d-e)). The density model

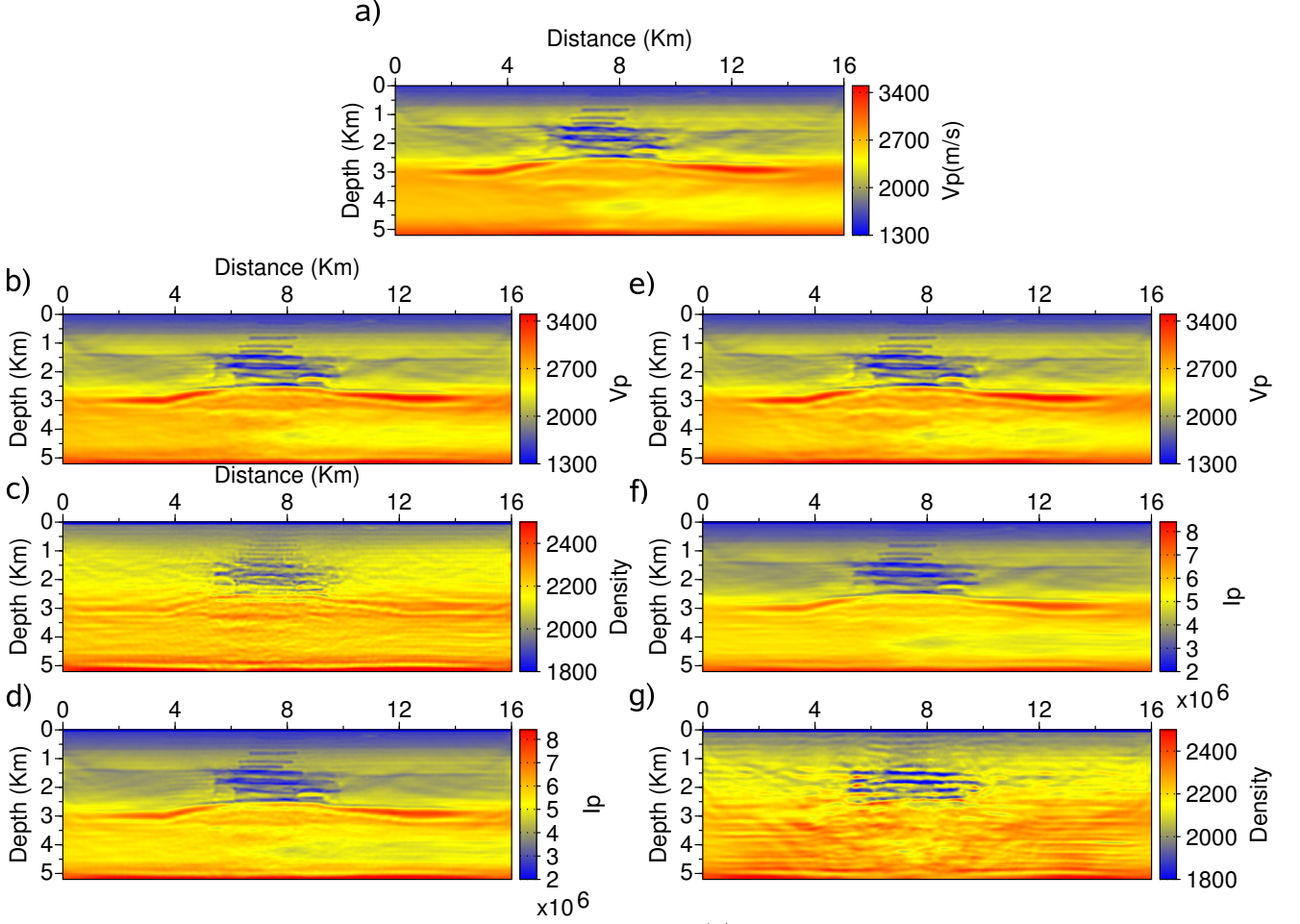


FIGURE 5.7 – Hierarchical update of two parameters. (a) P-wave velocity model inferred from a first mono-parameter inversion for the (V_P, ρ) parametrization, using a time damping of 1 s. (b-c) P-wave velocity and density models recovered by FWI for the (V_P, ρ) parametrization, and (c) impedance model derived from those two models. (d-e) P-wave velocity and impedance models found by FWI for the (V_P, I_P) parametrization, and (f) density model derived from those models.

derived from the two models is less unstable than the density model inferred from the simultaneous strategy with the (V_P, I_P) parametrization, but it is less accurate than the density models inferred from (V_P, ρ) parametrization (in one or two steps, as confirmed by the density errors in Table 5.2).

The model errors from the (V_P, I_P) parametrization are more important than from the (V_P, ρ) parametrization for each class of parameter, but we notice the mean time residual is the same for the two parametrizations. This highlights the non-linearity of the multiparameter inversion, as several model combinations can provide an equivalent match of the data.

By a way of conclusion, considering this surface acquisition and this case study, all the tests tend to show that the (V_P, ρ) parametrization is more suitable to recover both the velocity, the density *and* the impedance. This contradicts the generally accepted idea, which promotes

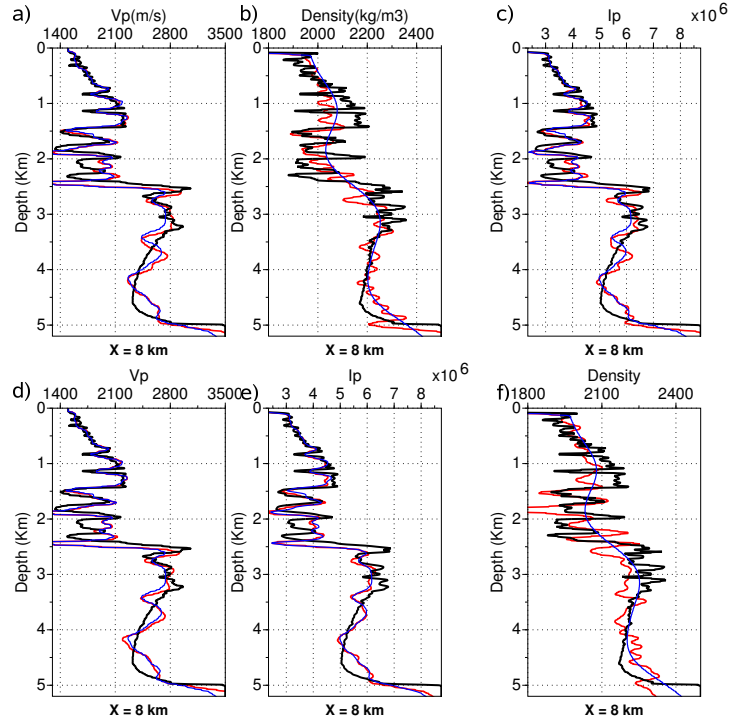


FIGURE 5.8 – Hierarchical update of two parameters. Logs extracted from the FWI models inferred from the second step of the hierarchical inversion. (a-b) P-wave velocity and density logs inferred by FWI for the (V_P, ρ) parametrization; (d-e) P-wave velocity and impedance models inferred by FWI for the (V_P, I_P) parametrization. (c,f) Impedance and density profiles inferred from the logs in (a,b) and (d,e) respectively.

the reconstruction of velocity and impedance. One reason is the more limited bandwidth reconstruction of the velocity in the (V_P, I_P) parametrization, which may impact the impedance reconstruction. A second reason is the small range of variation of the density relative to velocity and impedance, which prevents a stable reconstruction of density from a velocity and impedance model. This is why we will only consider in the following the (V_P, ρ) parametrization. Our study promote the use of the hierarchical strategy because the reconstruction of the density is more stable with it, and we noticed the model and time residuals are smaller than with the simultaneous inversion.

5.6 Reconstruction of velocity, density and Q_P -factor

In the following of this study, we seek to reconstruct the Q_P -factor in addition to the compressional wave speed and density. We first tune the full waveform inversion (FWI) regularization with the synthetic Valhall example before addressing the real data case study.

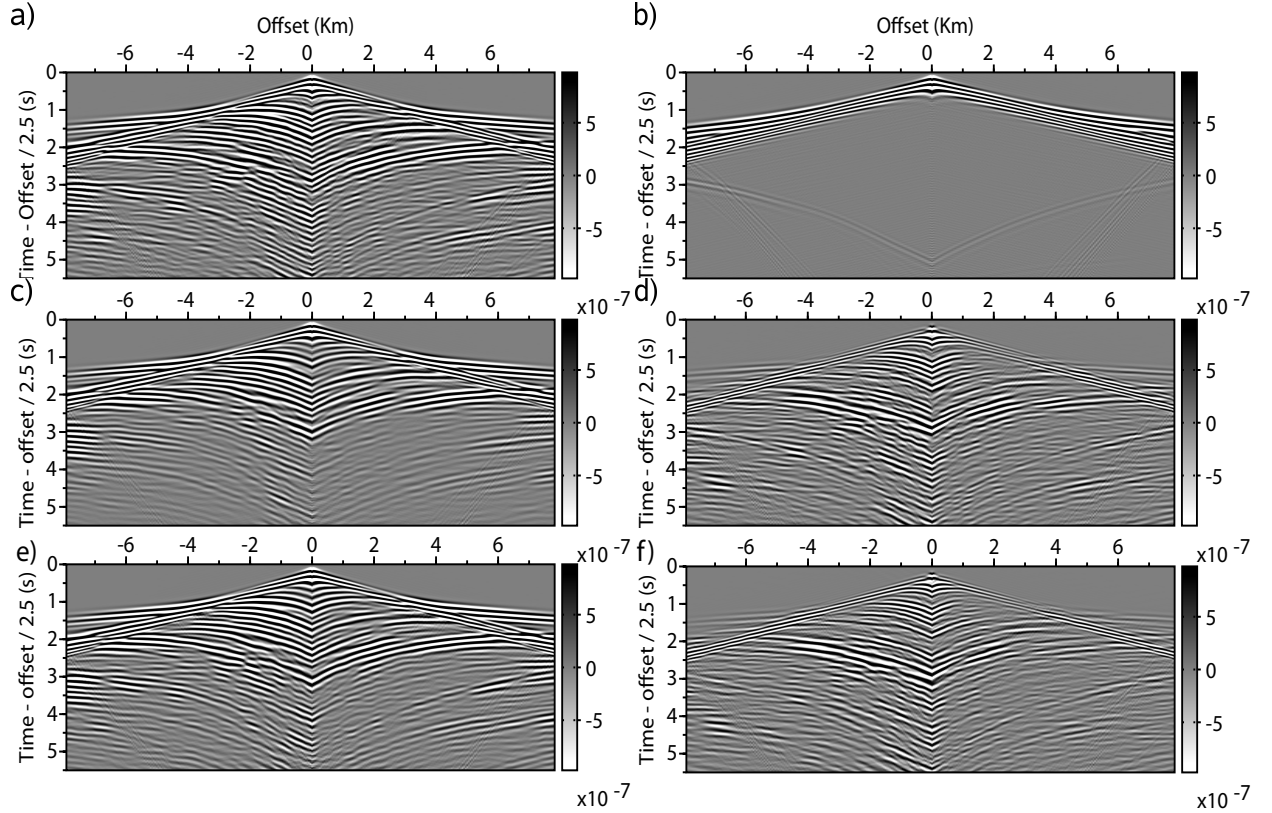


FIGURE 5.9 – Synthetic seismograms computed in (a) the true V_P , ρ and Q_P ($= 1000$) models; (b) the starting models of the first inversion step (smoothed version of V_P and ρ); and in the models obtained (c) at the first step of the hierarchical inversion when we only update the V_P parameter (Figure 5.7a); (e) at the second step of the hierarchical strategy for the (V_P, ρ) parametrization (Figure 5.7(b-d)). (d,f) Residual seismograms between the observed data (in a) and the seismograms shown in (c,e) respectively.

5.6.1 Synthetic Valhall case study

5.6.1.1 Sensitivity of data and monoparameter inversion to the density and Q_P -factor starting models

To mimic the experimental setup that will be used for the real data case study from Valhall, we consider VTI media, where the Thomsen's parameters were smoothed with a correlation length of 500 m and kept fixed during the inversion. We invert the same frequencies than in the previous section. The initial vertical velocity model is built by smoothing the true model with a Gaussian filter, where the correlation length is 250 m at the sea level, and 500 m at the bottom of the model. This should roughly mimic the resolution of traveltime tomographic methods. The density model is inferred from this velocity starting model using a Gardner law. The true and the starting models of the Q_P -factor are those presented in Figure 5.4(d,h,i).

Before showing the results of the multiparameter FWI, it is instructive to show how the inaccuracies of the velocity, density and Q_P -factor starting models translate into the data. For this purpose, we show in Figure 5.10(b-d) the difference between the seismograms computed in the true (V_{P0}, ρ, Q_P) models and the ones computed in two of the true models and in one initial

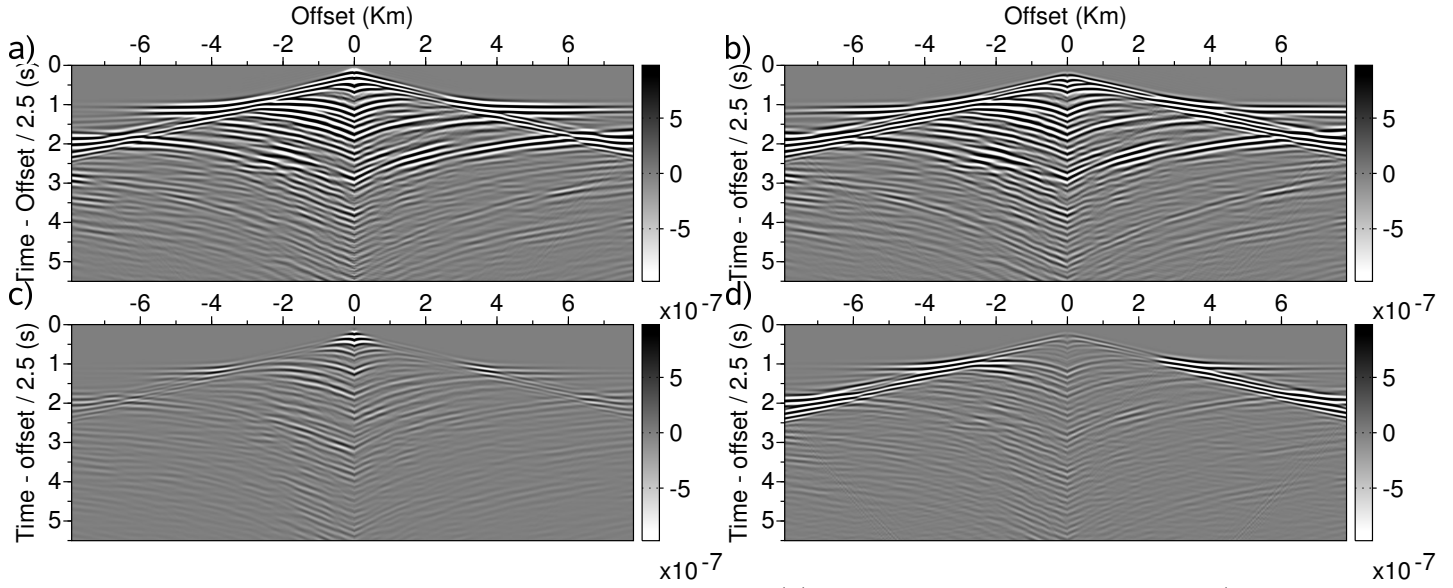


FIGURE 5.10 – Synthetic seismograms computed in (a) the true V_P , ρ and Q_P models (Figure 5.4(a,b,d)). (b-d) Seismograms corresponding to the difference between seismograms in (a) and the seismograms (not shown here) computed within the true models except (b) for the velocity model that is the smoothed version of the true model using a differential smoothing with depth, (c) for the density that is derived from the velocity starting model considered in (b) using a Gardner law, and (d) for the Q_P -factor model set to 150 in the earth and 1000 in the water (Figure 5.4h).

model. A smooth density model mainly impacts the match of the short-offset reflected waves, as a density diffractor leads to significant scattering at short apertures only (Figure 5.10c). The footprint of the homogeneous Q_P -factor model is visible in the residual seismograms over the full offset range, although it is stronger at large offsets as the waveform amplitudes decrease due to geometrical spreading, that is consistent with the isotropic radiation pattern of the Q_P -factor parameter (Figure 5.10d) (Malinowski *et al.*, 2011). The velocity smoothing leads to much stronger residuals relative to density and Q_P -factor, hence illustrating the dominant imprint of this parameter in the data (Figure 5.10b). All these statements are also shown by the time errors computed in the corresponding seismograms (Table 5.3), with an error for the velocity between 2 and 5 times higher than for the two other parameters. Time errors for the Q_P -factor are 2 times higher than for the density.

To assess the sensitivity of the monoparameter inversion for the P-wave velocity to the inaccuracies of the density and Q_P -factor background models, we compare the FWI velocity models inferred without time damping, when the background density and Q_P -factor models are the true and the smoothed models, respectively (Figures 5.11 and 5.12). The FWI models inferred from the smooth density and Q_P -factor background model are slightly noisier in the tertiary sediments, but remain close to the velocity model inferred from the true density and Q_P -factor background models. This validates the hierarchical approach, where the P-wave velocity is updated during a first inversion step. Percentage of error on the models is of 4.94 % and 4.77 % without and with the true models, respectively.

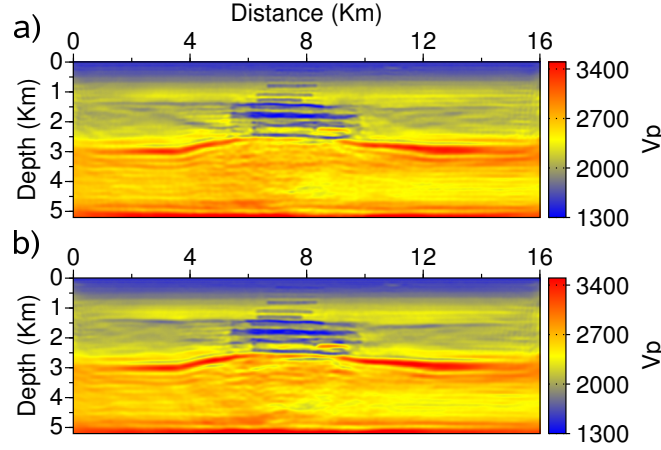


FIGURE 5.11 – Sensitivity of monoparameter FWI for V_P to the accuracy of the density and Q_P -factor background models. (a-b) P-wave velocity model inferred from monoparameter FWI using 3 time dampings. (a) The background density and Q_P models are smoothed version of the true models. (b) The background density and Q_P models are the true models.

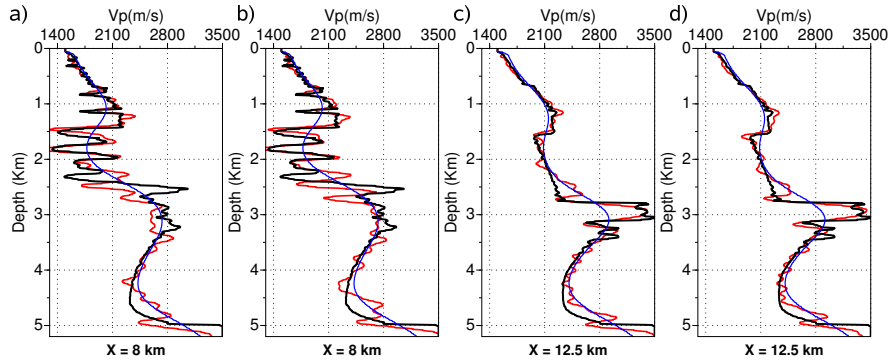


FIGURE 5.12 – Comparison between the true and FWI velocity profiles at $X = 8$ km (a-b) and $X = 12.5$ km (c-d) of distance. (a, c) The FWI logs are taken from the FWI model of Figure 5.11a. (b, d) The FWI logs are taken from the FWI model of Figure 5.11b.

5.6.1.2 Inversion strategies

We test three different inversion strategies to update the velocity, the density and the Q_P -factor. *Strategy 1* consists of the simultaneous update of the P-wave velocity, density and Q_P -factor without time damping.

The simultaneous inversion should prevent artifacts, which arise when hierarchical strategies are used. In this setup, the risk is that the dominant parameter can contribute to reduce residuals mainly associated with secondary parameters. The *second strategy* is a multiscale procedure, which proceeds hierarchically over the data and the parameter classes. In a first step, we jointly update the large wavelengths of P-wave velocity and Q_P -factor using an aggressive time-damping factor τ of 1 s, equation 5.8, in order to minimize data residuals associated with

the wide aperture angles before the simultaneous update of the shorter wavelengths of P-wave velocity, density and Q_P -factor without time damping to reduce the remaining residuals at short apertures. Following a multiscale approach, low wavenumbers are reconstructed first by updating the parameters V_P and Q_P , which are sensitive to the wide apertures, before the reconstruction of the short wavelengths by updating the parameters, which are sensitive to the short apertures (V_P , ρ and Q_P). In the *third strategy*, we update the P-wave velocity without time damping during a first step to cancel most of the residuals and build a kinematically-accurate velocity model before updating jointly the three parameter classes. Of note we keep the same starting model of density for the first and the second inversion steps.

Compared to the update of the density or impedance presented in the previous section, an additional difficulty with the reconstruction of the Q_P -factor results from the weak imprint of attenuation in the data for realistic marine targets (Malinowski *et al.*, 2011). We verified this statement for the Valhall case study where the diagonal terms of the Hessian (the auto-correlation of the partial derivative of the data with respect to each parameter class) have amplitudes of the order of 10^{-17} , 5.10^{-18} and 10^{-21} for V_P , ρ and Q_P , respectively. Applying the inverse of the Hessian to the gradient of the misfit function should theoretically account for the different sensitivity of the data to each parameter class. However, a difficulty arises when a damping regularization term λ_i , equation 5.3, is added to the diagonal of the Hessian for regularization. If the damping λ_i are too high, they can annihilate the scaling effect of the Hessian during the update of the weak parameter and also impacts the reconstruction of the other parameters. Therefore, we found that a suitable choice of the damping λ_i was a key feature to scale properly the model perturbations associated with each parameter classes.

To understand the effect of the damping terms λ_i on the inversion, we performed a series of tests, an outline of which is provided below. First, we estimate a suitable value of the damping term for the velocity (λ_{V_P}) during a mono-parameter inversion. As an illustrative example, a damping term of 10^{-6} leads to a model gradient ten times higher than the data gradient (see the right terms on Equation 5.3 for a definition of the so-called data-space and model-space gradients), resulting in negligible velocity perturbations. A good final model of P-wave velocity is obtained using a damping term of 4.10^{-8} , leading to a model gradient ten times smaller than the data gradient. Second, we perform the simultaneous updates of P-wave velocity, density and Q_P -factor using the same damping for each parameter class (4.10^{-8}). This inversion setup leads to underestimated model perturbations, that prompt us to repeat the experiment with decreasing values of the damping until unstable reconstructions are found. The smallest value of the damping, which leads to reasonable models, was found equal to 4.10^{-18} , that is ten times smaller than the maximum coefficient of the Hessian block associated with the velocity parameter (Figure 5.13). Decreasing the damping to a value as small as 8.10^{-20} for the three parameters leads to an unperturbed velocity model, a contrasted model of density and an unstable model of the Q_P -factor. We conclude from these experiments that a suitable approach to scale properly the model perturbations associated with each parameter class consists of adapting the value of the damping factor to the parameter class. We use a value of the damping term for Q_P , that is 50 times smaller than the dampings used for the P-wave velocity and density. It is worth noting that, even in this setting, the damping term of the Q_P -factor is almost three orders of magnitude higher than the maximum coefficient of the diagonal block of the Hessian associated with the Q_P -factor, hence, making negligible the contribution of this part of the Hessian.

5.6.1.3 FWI results

We synthesize all of the results (models and logs) in Figures 5.13 and 5.14, and the corresponding time and model errors ($\xi(\mathbf{m})$ and $\xi(\mathbf{s})$) are shown in Table 5.3 to appraise them. The FWI Q_P models can be compared to the true model represented with the appropriate color scale in Figure 5.4i. We show the inversion results for the three strategies described in the previous section for two sets of dampings λ_i during the joint update of P-wave velocity, density and Q_P -factor : ($\lambda_{V_P} = 4.10^{-18}$, $\lambda_\rho = 4.10^{-18}$, $\lambda_{Q_P} = 8.10^{-20}$) and ($\lambda_{V_P} = \lambda_\rho = \lambda_{Q_P} = 4.10^{-18}$). A first overview over the whole scene shows large variation in the solutions, in particular for the density and the Q_P -factor, that reminds us the intrinsic non linearity of the multiparameter inversion. It is worth noting that the value of λ_{Q_P} not only impacts the reconstruction of the Q_P -factor parameter, but also all of the other parameters.

Strategy 1 : joint update of multiple parameters

For the *first strategy*, the time and model errors (except for the velocity) tend to promote the use of a small value of λ_{Q_P} (i.e., 8.10^{-20}). With this setting (Figure 5.13(a-c)), the density and the Q_P -factor perturbations are better balanced compared to the case where the same damping is used for all of the parameters (Figure 5.13(d-f)). In this case, the Q_P -factor perturbations are clearly underestimated, while the density perturbations are overestimated (compare Figure 5.13(a-f) and 5.14(a-f)). The drawback of this strategy is that the velocity model shows some inaccuracies at the reservoir level between 2.2 and 3 km in depth.

Strategies 2 and 3 : hierarchical approaches

In the strategies 2 and 3, we make some attempts to increase the quality of the starting models of the second-step inversion, the aim of which is to jointly update the 3 parameters. The *strategy 2* is intermediate between the strategies 1 and 3 : while *strategy 1* does not include a preliminary inversion step, the first inversion step of the *strategy 3* uses the full data space to update a reduced model space corresponding to the dominant parameter V_P , hence putting strong prior information in the initial V_P model of the second-step inversion. The *strategy 2* uses a reduced data space (the wide-aperture components of the data) to update a reduced model space involving two parameter classes during the first inversion step. Contrary to the *strategy 1*, time and model errors in the second and third strategies tend to promote the use of the same damping term for all of the parameters. However, we can question the relevance of the model error as it depends on the wavenumber content of the model perturbations.

The *strategy 2* succeeds in the reconstruction of smooth velocity and Q_P -factor model during the first step of the hierarchical approach (Figure 5.13(g-h)). This helps to build a more accurate velocity model during the second step of the *strategy 2* than during the *strategy 1* : although a mispositioning in depth of the top of the reservoir subsists, the velocities are better reconstructed in terms of amplitude. However, the reconstruction of the density and the Q_P -factor is not successful during the second-step and shows noisy reconstructions in particular for the density. The *strategy 3* (Figure 5.13(o-u)) leads to a final velocity model, which is close to the one inferred from the *strategy 2*. A more stable reconstruction of the Q_P -factor and density is however achieved because most of the residuals have been canceled out during the update of the dominant parameter during the first-inversion step (Figure 5.13o). The underestimated perturbations in the density and Q_P -factor models highlight, however, the trade-off between velocity and these two secondary parameters. The trade-off between the two secondary parameters, ρ and Q_P , is highlighted by the results obtained with the two different values of λ_{Q_P} :

as λ_{Q_P} is decreased the Q_P -factor perturbations are strengthened at the expense of the density perturbations.

Summary of the synthetic example results

This study has highlighted the difficult problem of FWI tuning when multiple classes of parameters must be updated with potential trade-off between parameters and variable sensitivity of the data to the parameters. The choice of the best damping factor in the Hessian has been shown to be a key issue to guarantee reliable results. Our approach to estimate these dampings (one per parameter class) remains heuristic and relies on trial-and-error approaches. One may want to use the L-curve method to determine the best damping terms (Hansen, 1992), but it would be very time consuming to do it for each of the three parameters. Despite our indicators of quality, it is difficult to decide which model is the most satisfying. For real case applications, our geological a priori and the a priori information coming from well logs should help to choose the best parametrization. One may also want to adopt the famous principle : *When in doubt, smooth* (Sir Harold Jeffreys, quoted by Moritz & Sunkel (1978)), also known as the Occam's principle. A Bayesian view to smooth the models is to take the mean over all the derived FWI models. We applied this idea to the models presented in Figure 5.13, leading to the mean models shown in Figure 5.15. It is interesting to see how this process decreased the noise that was present in the original models. We see that the mean Q_P -factor allows us to accurately locate in depth the gas layers, although the amplitude is still underestimated. The mean error on those models (0.87) is equal to the smallest mean error obtained for the models of Test 8, and the time error is the second smaller value among all the other tests. Of note we took a simple mean of the models, but one could compute a weighted mean, where the weights would be defined from the model and time errors to enhance the most realistic models (only possible for the synthetic case). For real cases, a priori criteria like in-situ petrophysical measurements can be used to define probability densities via Bayes Theorem (Gouveia & Scales, 1997, 1998).

Test No.	STEP 1	Test STEP 2	$\xi(\mathbf{m})$			$\xi_{mean,n}(\mathbf{m})$	$\xi(\mathbf{s}_1)$	$\xi(\mathbf{s}_2)$	$\xi(\mathbf{s}_3)$	$\xi_{mean}(\mathbf{s})$	Q_P	λ	ρ
			V_P	ρ	Q_P		$\cdot 10^{-5}$	$\cdot 10^{-5}$	$\cdot 10^{-5}$	$\cdot 10^{-5}$		V_P	
-		$V_{P,0}$	5.09	0	0		1.73	3.30	1.85	2.29			
-		ρ_0	0	2.26	0		0.31	0.65	0.31	0.43			
-		$Q_{P,0}$	0	0	32.66		0.83	1.2	0.82	0.95			
1		$(V_P, \rho, Q_P)_{no-dp}$	5.03	2.53	25.3	0.91	1.33	2.26	1.3	1.63	$8 \cdot 10^{-20}$	$4 \cdot 10^{-18}$	$4 \cdot 10^{-18}$
2		$(V_P, \rho, Q_P)_{no-dp}$	4.90	2.82	27.86	0.96	1.53	2.45	1.50	1.83	$4 \cdot 10^{-18}$		
3	$(V_P, Q_P)_{1dp}$		5.20		28.89		1.51	2.67	1.63	1.94	$8 \cdot 10^{-20}$		
4	$(V_P, Q_P)_{1dp}$	$(V_P, \rho, Q_P)_{no-dp}$	5.09	2.96	27.84	0.98	1.07	2.00	1.07	1.38	$8 \cdot 10^{-20}$		
5	$(V_P, Q_P)_{1dp}$	$(V_P, \rho, Q_P)_{no-dp}$	5.12	3.04	24.65	0.96	1.02	2.07	0.97	1.35	$4 \cdot 10^{-18}$		
6	$(V_P)_{no-dp}$		4.94				1.43	2.25	1.41	1.70	$8 \cdot 10^{-20}$		
7	$(V_P)_{no-dp}$	$(V_P, \rho, Q_P)_{no-dp}$	4.89	2.41	26.11	0.89	1.09	1.79	1.18	1.35	$8 \cdot 10^{-20}$		
8	$(V_P)_{no-dp}$	$(V_P, \rho, Q_P)_{no-dp}$	5.00	2.39	23.54	0.87	0.89	1.68	0.91	1.16	$4 \cdot 10^{-18}$		
-		Mean FWI models	4.72	2.41	25.08	0.87	1.00	1.70	1.05	1.25			

TABLE 5.3 – Table referencing time and model errors computed within the different FWI models presented in Figure 5.13. We want to compare the three following strategies : 1) inversion of V_P , ρ and Q_P in one step without time damping $[(V_P, \rho, Q_P)_{no-dp}]$; 2) inversion of V_P and Q_P at the first step with a time damping of 1 s $[(V_P, Q_P)_{1dp}]$, followed in a second step by the simultaneous inversion of V_P , ρ and Q_P $[(V_P, \rho, Q_P)_{no-dp}]$; 3) inversion of V_P first without time damping $[(V_P)_{no-dp}]$, followed by the simultaneous inversion of V_P , ρ and Q_P without time damping.

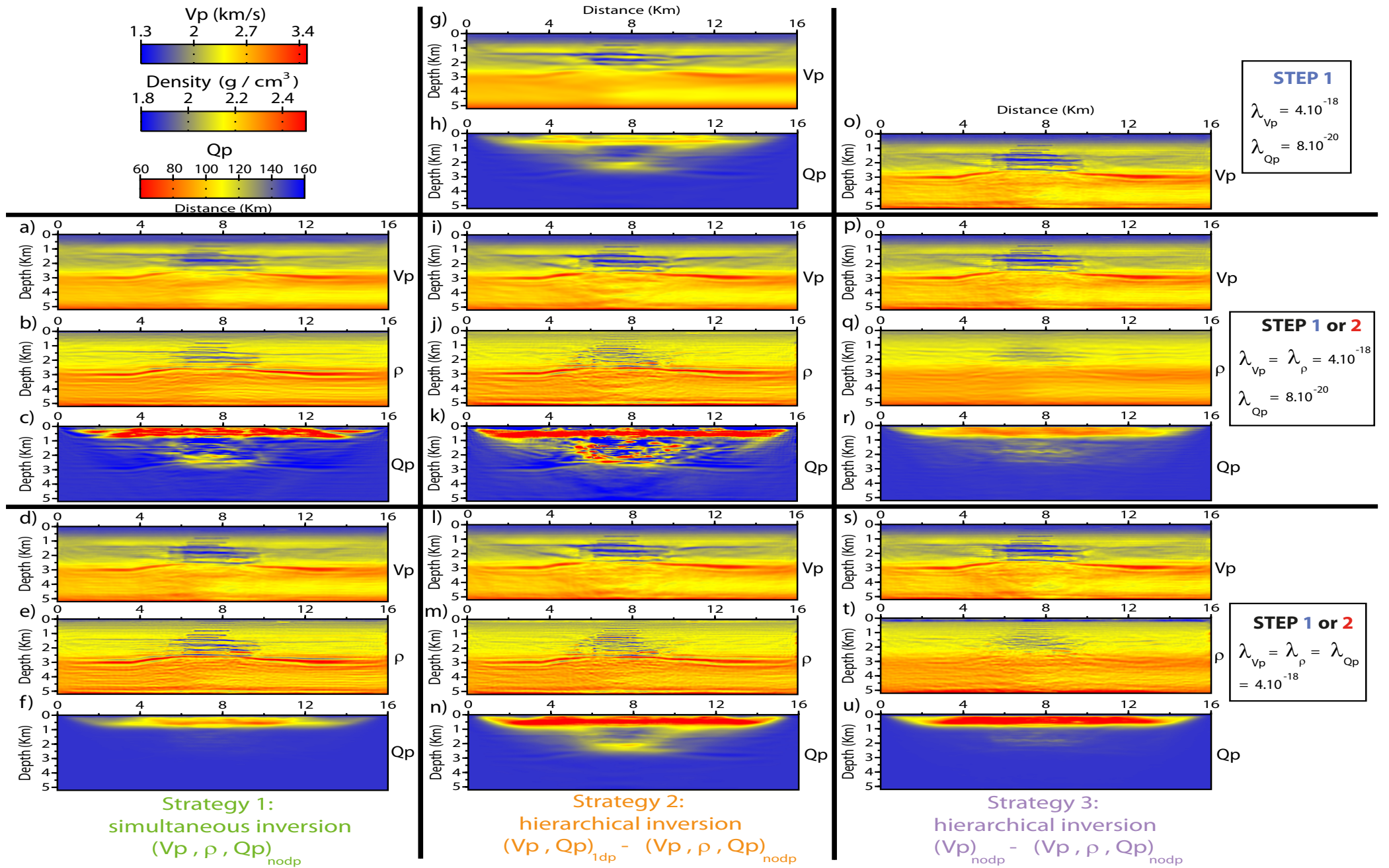


FIGURE 5.13 – Imaging V_P , ρ and Q_P by FWI. (a-u) Velocity, density and Q_P -factor models inferred from three different inversion strategies (see text for details). For each strategy, the FWI was performed using the same damping ($= 4.10^{-18}$) for all of the parameter classes, and using λ_{Q_P} 50 time smaller than λ_{V_P} and λ_{ρ} .

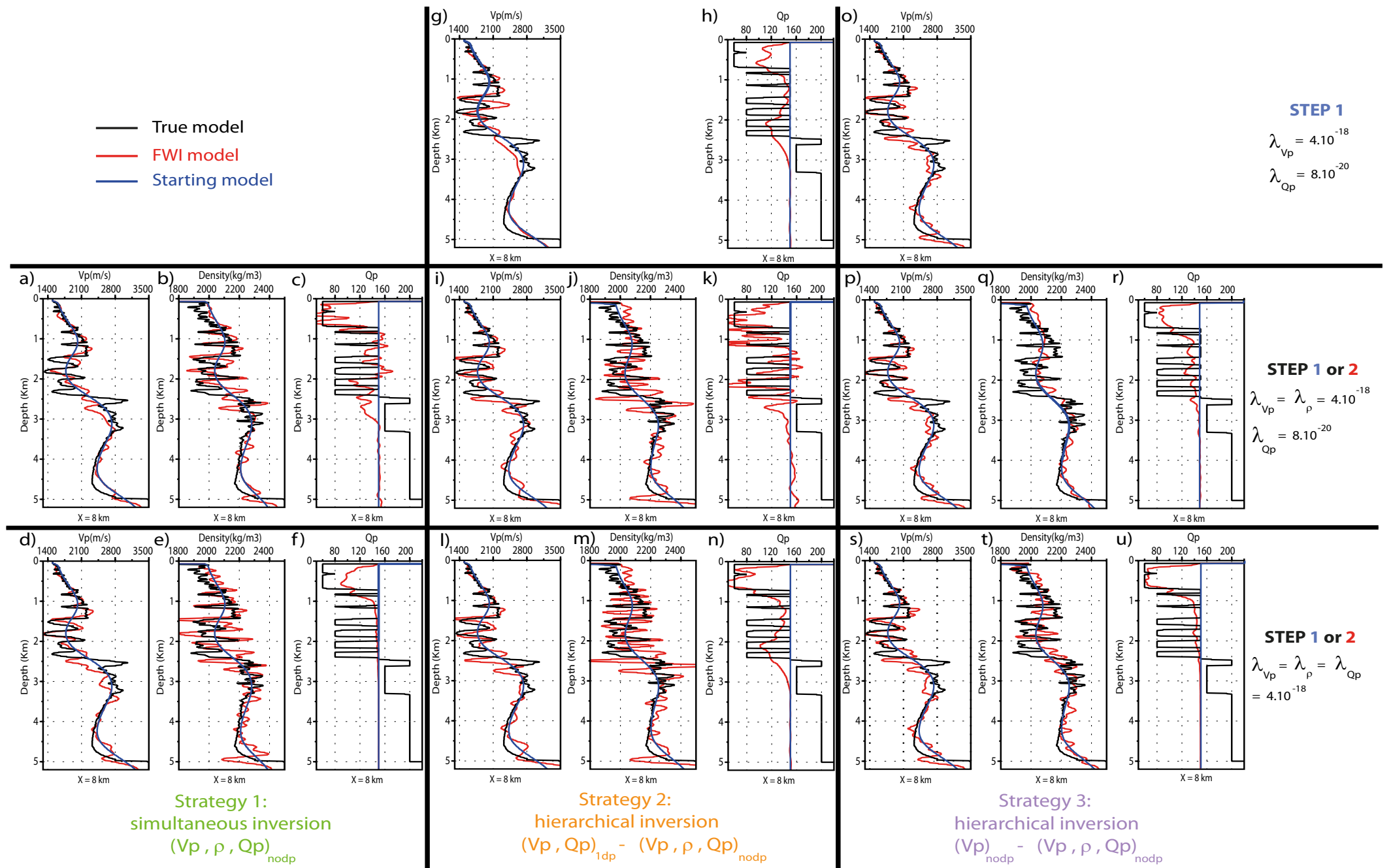


FIGURE 5.14 – (a-u) Velocity, density and Q_p -factor profiles extracted from the models shown in Figure 5.13(a-u).

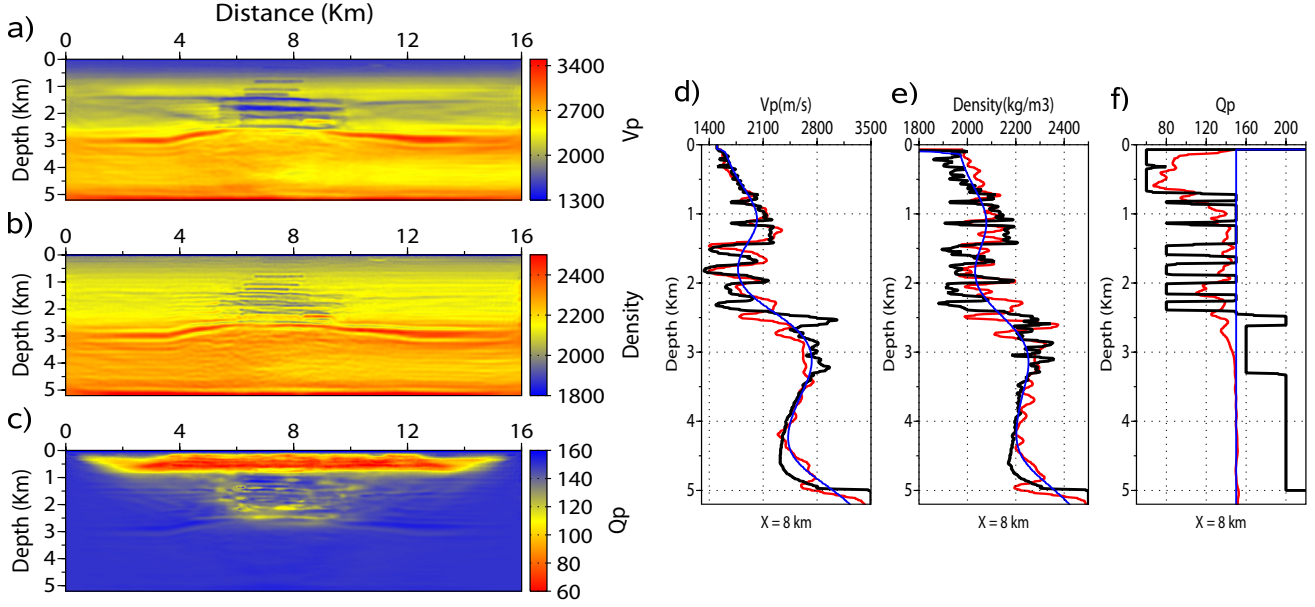


FIGURE 5.15 – Mean FWI models (a-c) and their vertical profiles (d-f). The mean models are obtained by averaging all of the models of Figure 5.13. (a, d) Mean V_P model. (b, e) Mean density model. (c, f) Mean Q_P -factor model.

5.6.2 Application to real OBC data from the Valhall field

We now present the application of multiparameter anisotropic acoustic full waveform inversion (FWI) to two cables (21 and 29) of a 4C OBC survey from the Valhall field, which are positioned on the survey map in Figure 5.16. Acoustic isotropic FWI of this data set was already presented by Sirgue *et al.* (2010). The resulting velocity model shows a complex network of channels 150 m below the surface as well as a gas cloud with gas-filled fractures at 1 km in depth.

The hydrophone data of cable 21 was processed by anisotropic and isotropic acoustic FWI by Prieux *et al.* (2011) to assess the footprint of anisotropy on isotropic FWI. They show that the horizontal velocities are reconstructed by isotropic FWI in the upper part of the target, where the FWI is mainly driven by the wide aperture components of the data (diving waves and super-critical reflections). This can lead to underestimated velocities in the gas layers and/or mispositioning of the reflectors at the reservoir level, that are required to match the short-spread reflections from the deep discontinuities. In the following, we shall take into account anisotropy by using large-scale models of the Thomsen's parameters, that were already built by anisotropic reflection traveltime tomography. Prieux *et al.* (2011) have shown that, for the Valhall case study, the long wavelengths of the Thomsen's parameters provide sufficiently accurate background models for a reliable reconstruction of the vertical velocity by FWI. The shot and receiver spacings are 50 m and the maximum offset in the data is 14.5 km. A more detailed description of the anatomy of the Valhall OBC data is provided in Prieux *et al.* (2011).

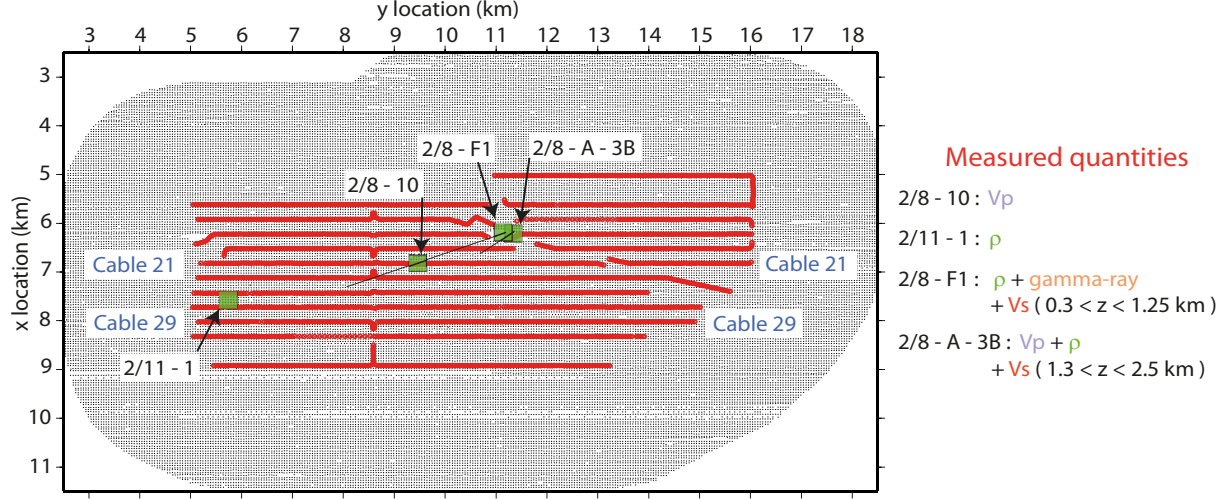


FIGURE 5.16 – Valhall acquisition layout on which we added location of the well logs provided by BP (green squares) measuring different types of quantities (indicated on the right). Of note, the wells '2/8-F1' and '2/8-A-3B' are in the close vicinity of the Valhall platform. We added from their location a black line denoting their route in depth.

5.6.2.1 Inversion setup and starting models

Only the hydrophone component is considered for FWI. We invert the five overlapping groups of three frequencies : [3.5, 3.78, 4], [4, 4.3, 4.76], [4.76, 5, 5.25], [5.25, 5.6, 6] and [6, 6.35, 6.7] Hz. We test the *strategies 1* and *3*, that were assessed during the previous synthetic case study.

The anisotropic visco-acoustic wave equation is parametrized by the vertical velocity, the density, the quality factor and the Thomsen's parameters, and we keep fixed the anisotropic parameters during the inversion (Prioux *et al.*, 2011). The initial density model of one inversion step is inferred from the initial velocity model of the current inversion step by the Gardner law. We show the starting models for the velocity, the Thomsen's parameters and the density of line 21 and line 29 in Prioux *et al.* (2011) (their figure 2), and in Figure 5.17.

In order to build a realistic Q_P -factor starting model, we computed visco-acoustic seismograms considering the V_{P0} starting model of line 21 (Prioux *et al.* (2011), their figure 2a), and we test different homogeneous values of Q_P below the sea bottom to match the AVO trend of the first arrivals. We focused on the first 0.5 s of the signal after the first arrival, and we computed the L2 norm of the windowed signal for each offset (Figure 5.18). A value of 150 for Q_P gives in average the best agreement with the main trend of the recorded AVO curve.

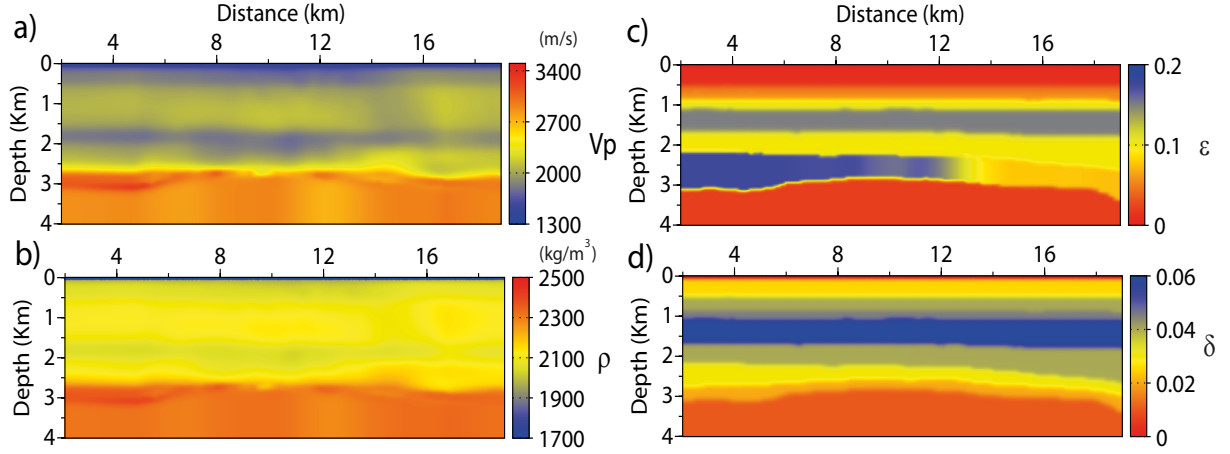


FIGURE 5.17 – Starting models of line 29 for the vertical velocity (V_{P0}) (a), the density (b) derived from V_{P0} model using a Gardner law, and for the Thomsen's parameters ϵ and δ (c-d).

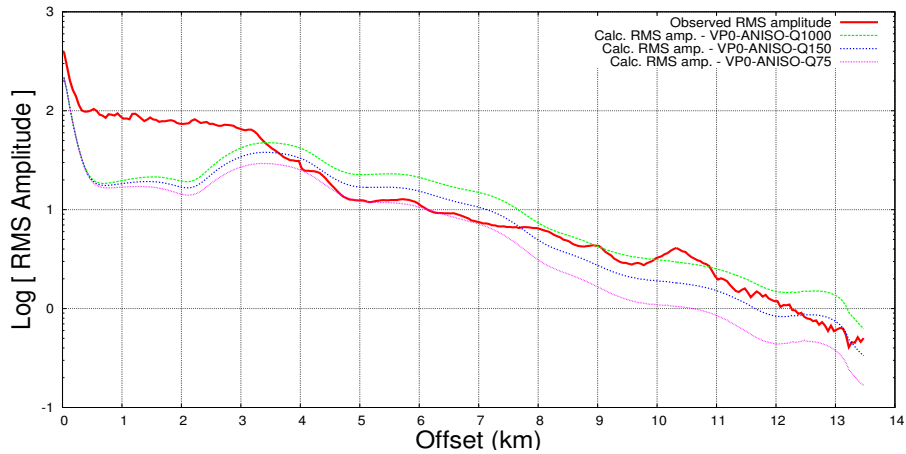


FIGURE 5.18 – AVO trend of the first-arrival travel times plotted in a logarithmic scale, for the observed data (in red), and the computed data using an homogeneous value of Q_P equal to 1000 (in green), 150 (in blue) and 75 (in pink). The best match is obtained for a value of Q_P equal to 150.

5.6.2.2 FWI results and model appraisal

The final FWI models of lines 29 and 21 inferred from *strategies 1* (joint update) and *3* (hierarchical update) are shown in Figure 5.19 and 5.20 respectively. Following the results of the synthetic experiment, two combinations of dampings λ_i are tested : the same damping is used for all parameters (Figure 5.19(a-c,h-j)) or the damping term of Q_P is 100 times smaller than the dampings for velocity and density (Figure 5.19(d-f,k-m)). We do not show the results obtained with the *strategy 2* of the synthetic case because the perturbations for the three parameters are too small.

We follow the same heuristic than for the synthetic experiment to define suitable values

of the dampings λ_i . Compared to the synthetic case, we notice that the range of acceptable values of λ_{V_P} during mono-parameter inversion is much more narrow than for the synthetic experiment. This different behavior is probably due to the noise in the real data. It follows that the contribution of the diagonal block of the Hessian for each parameter is negligible relatively to the damping terms. The data Hessian for the velocity, the density and the Q_P -factor vary within the same proportion than for the synthetic case study.

When the same damping is used for all of the parameters, the Q_P -factor model is left almost unchanged when *strategies 1* and *3* are used (Figure 5.19(d-f,k-m)). Moreover, the velocity model is noisier than the one obtained with a small value of λ_{Q_P} . Therefore, we will focus in the following on the interpretation of the results obtained with a small value of λ_{Q_P} .

FWI allows us to image several density contrasts along lines 21 and 29 (see one of the black arrow in Figure 5.19b). In particular, we show a deep reflector at 3.5 km in depth bending from the middle of the model to the bottom left, that was already imaged by reverse time migration (Prioux *et al.*, 2011). The low density contrasts at 0.6 km and 1.5 km in depths are interpreted as an accumulation of gas below barriers of claystones. Note that this density contrasts are more pronounced for line 29 at around 7 km and 11 km of distance, just above two low velocity areas probably associated with gas-charged sediments (white dashed circles in Figure 5.19h). The FWI velocity model inferred from *strategy 1* is smoother than the one inferred from the *strategy 3*, that is balanced by more contrasted density and Q_P -factor models when the *strategy 1* is used. This is consistent with the overall trend revealed by the synthetic experiment and highlights the trade-off between parameters. The velocity models obtained close of the first-step and second-step inversions are almost the same when *strategy 3* is used, which highlights the dominant weight of the velocity in the inversion.

The Q_P -factor models show interesting features amenable to geological interpretations. We show weak values of Q_P down to 600 m in depth in both models of lines 21 and 29. We interpret these attenuating zones as the signature of soft quaternary sediments. Small values of Q_P (highlighted by white circles in Figure 5.19(c)) are also shown in the model of line 29 at 8 km and 11 km of distances and at 2.5 km in depth, that are the same locations where we observed smaller values of V_P . This zones are located at the bottom of two vertical low- Q_P features that may be interpreted as gas chimney (Figure 5.19j). Line 21 is known to be located above tertiary sediments strongly charged in gas (Sirgue *et al.*, 2010), as confirmed by the final velocity models of Figure 5.20(a,e). Consequently, the Q_P factor present smaller values in the middle of the model of line 21 compared to line 29. Another interesting feature is visible only in the Q_P -factor models of line 21 at [11.5-12] km of distance and 1 km in depth bending on the bottom right (white circles in Figures 5.20(c,g)).

This feature may be interpreted as a gas-filled fracture, like the one imaged in Sirgue *et al.* (2010) (their figure 4c) although we reach a lesser resolution than in 3D-FWI. This feature is imaged with a stronger amplitude but with less resolution when the simultaneous inversion is used relatively to the hierarchical inversion, as already reported by (Kamei & Pratt, 2008).

The quality factor is much lower in the first 600 m depth than in the gas, despite we know the intrinsic attenuation of the gas is very important. In addition to the loss of resolution with depth and the lithological differences, this observation may be related to the fact the inverse of the quality factor is strongly dependent on effective pressure, decreasing by at least one order of magnitude between ambient pressure and 40 MPa, corresponding to a depth burial of about 2000 m (studies made on sandstones in Winkler & Nur (1982)).

In order to enhance the information on the reflectivity provided by the reconstruction of the density, we show the density perturbations for the two lines 21 and 29 and for the two strategies (Figure 5.21). We note an overall good lateral continuity of the reflectors. The simultaneous inversion mainly enhances two reflectors at 2.6 km and 1.5 km in depth. Amplitude of these perturbations are higher than the amplitude of the perturbations built by the second step alone of the hierarchical inversion (Figure 5.21(c,f)), especially for line 21, and are of the same order of magnitude than the total perturbations of the hierarchical inversion (also seen in Figure 5.22(e-h)). We conclude that the density starting model built with a Gardner law for the second inversion step of the hierarchical approach contains most of the high-wavenumber perturbations of the final density model. Although this strategy allows us to inject prior information on the density from the Gardner law, it may in the other hand drive the convergence towards a local minimum. We also notice the density contrast associated with the deep reflector at 3.5 km in depth (black arrows in Figure 5.21b), already shown in the FWI velocity model. The weaker perturbations of line 21 compared to line 29 at the second inversion step can be related to the platform noise that could hamper the convergence, or by the stronger presence of gas for line 21. Interestingly, we notice in the residuals of the second step for line 29 a lateral discontinuity below 2.5 km depth (black ellipse) that may indicate the presence of a fault. This discontinuity is also present in the total density residual.

Comparison with sonic logs

To further appraise our results, we compare FWI profiles with corresponding band-pass filtered well logs in the [0-14] Hz frequency range, the location of which is indicated in Figure 5.16. The match between FWI and well logs should be assessed taking into account that most of those logs are not located on lines 21 and 29. Moreover, the logs are not rigorously vertical, as shown for the '2/8-F1' and '2/8-A-3B' wells by the black line starting from the green square (Figure 5.16).

For velocity, we show a good agreement at 750 m in depth with the '2/8-10' sonic log, and a reasonable match below, especially for line 21 located at the same cross-line location than the well (Figure 5.22a). The match with the '2/8-A-3B' log is less obvious, in particular for line 29 whose cross-line position is farther to the well.

A very strong low density contrast is present in the '2/11-1' and '2/8-A-3B' density well logs at about 1.5 km in depth. This contrast is detected by FWI but with an underestimated amplitude. It is worth noting that the FWI density logs show a deficit of small wavenumbers between 1.7 and 2.5 km in depth, that results from the lack of sensitivity of wide aperture data to density. We show also a good match between the well logs and the FWI logs in the upper part of the target down to 700 m in depth, in particular for the '2/8-F1' well log (Figure 5.22(f,h)).

The '2/8-F1' well performs gamma-ray logging, hence measuring the clay content of the subsurface. According to Klimentos & McCann (1990), the intra-pore clay content of sandstones sediments is closely related to the attenuation of the compressive velocity, and to the rocks permeability. The empirical law established between the clay content C (in percents) and the Q_P -factor Q_P is $Q_P = 179 C^{-0.843}$. We compute the attenuation derived from the measured clay content (Figure 5.22(i-k)), that we multiplied by a factor 10 for an easier comparison. Without this scaling the measured Q_P -factor is about 15, that seems an unrealistic value, and variations would be too small to be seen. It may be related to the fact the well was drilled in shale sediments, while the empirical relation was derived for sandstones. Despite this, we

can use it to assess attenuation variations in a relative sense. In the Q_P -factor profiles (Figure 5.22(i-k)), Q_P reaches a minimum value of 50, corresponding to the lower limit that we allowed for this parameter. This value was roughly chosen based on our a priori and the AVO curves shown in Figure 5.17. Of note, our approach to bound the values of the parameters is quite crude, and could be improved using an extension of the L-BFGS, called L-BFGS-B (last B for bounds) (Byrd *et al.*, 1995; Zhu *et al.*, 1997). The log and the FWI Q_P -factor profiles vary in the same direction between 250 m and 700 m in depth, and the comparison between the two profiles is more difficult below. For line 21, the large spatial wavelengths of the Q_P -factor profile obtained by hierarchical inversion tends to reproduce the very low wavenumbers of the well log (large ellipse between 1 km and 2.6 km depth in Figure 5.22i). Below 700 m in depth, we highlighted with some black dashed ellipses the different depths where the FWI profiles and the well log share common features.

Data match

We show the direct comparison between the recorded and the computed data within the final models obtained for the two OBC lines by simultaneous or hierarchical inversion in Figure 5.23. The source wavelet used for the seismic modeling was estimated in the V_{NMO} model built by reflection traveltime tomography from offsets smaller than 2 km (Prioux *et al.*, 2011). The long-offset refracted waves computed in the FWI model obtained after the second inversion step of the hierarchical approach are slightly in advance relatively to the recorded arrivals and amplitudes are weaker than that obtained after the first inversion step. Both effects are attributed to the low values of quality factor in the first 600 m in depth in which refracted waves propagate. The small changes of velocities between the first and second steps were not significant enough to correct for this small phase shift. In the contrary, the amplitude and phase match of the diving waves in the models obtained by simultaneous inversion are better for line 21 and for a line 29 to a smaller extent. This improved match might be permitted by the simultaneous inversion update of the three parameter classes, that gives more freedom to the inversion to match the data. In the other hand, seismograms from line 29 (only) indicate the second step inversion strengthen and improve the short-spread reflections, which also appear with a stronger amplitude than in the seismograms from the simultaneous inversion.

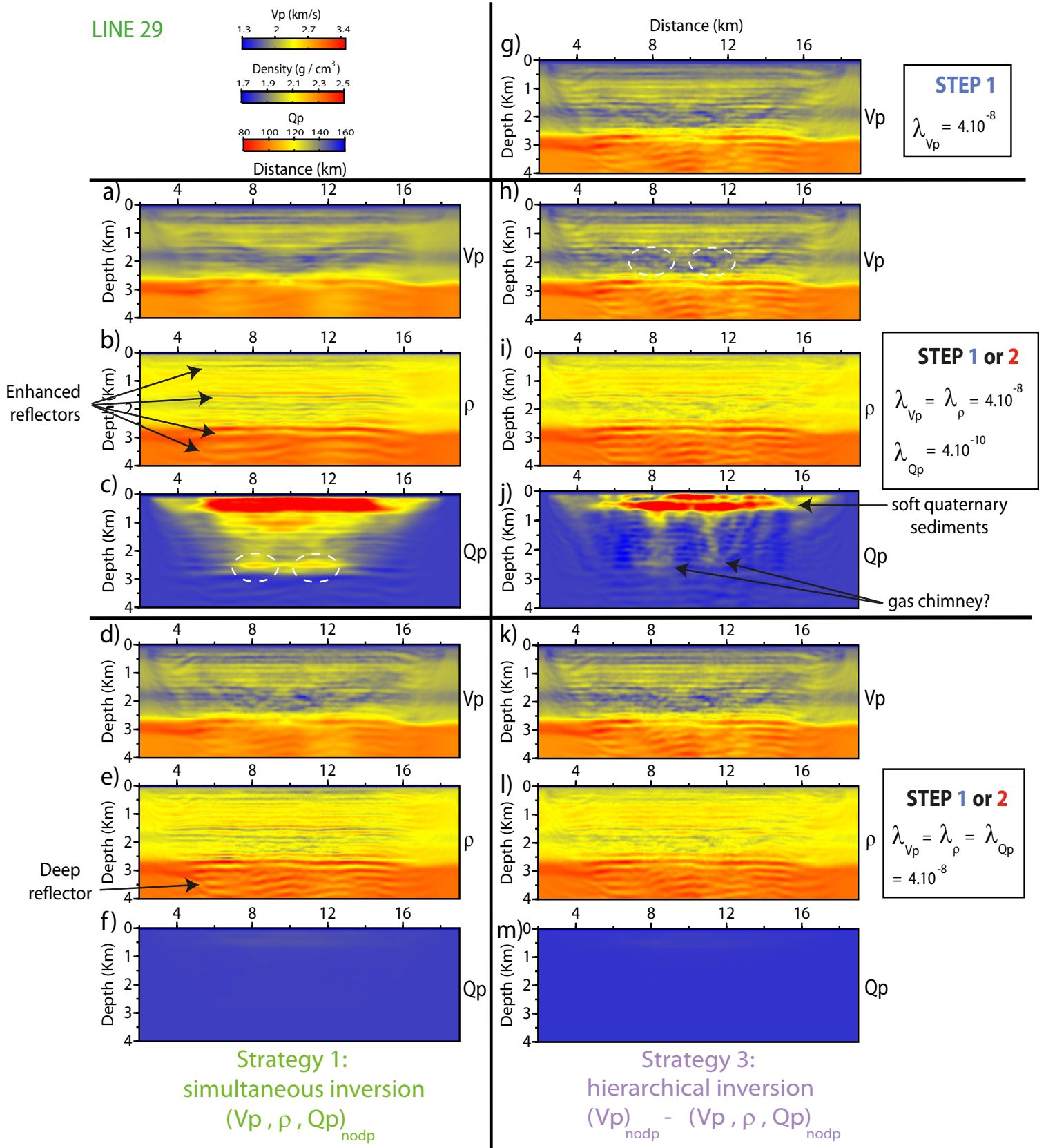


FIGURE 5.19 – FWI models for vertical velocity, density and Q_P -factor for line 29. (a-f) simultaneous update of multiple parameters. (g-m) Hierarchical update of multiple parameters. The damping terms are the same for all the the parameters in (d-f,k-m). The damping term associated with the Q_P -factor parameter is 100 times smaller than those used for the other parameters (a-c,h-j).

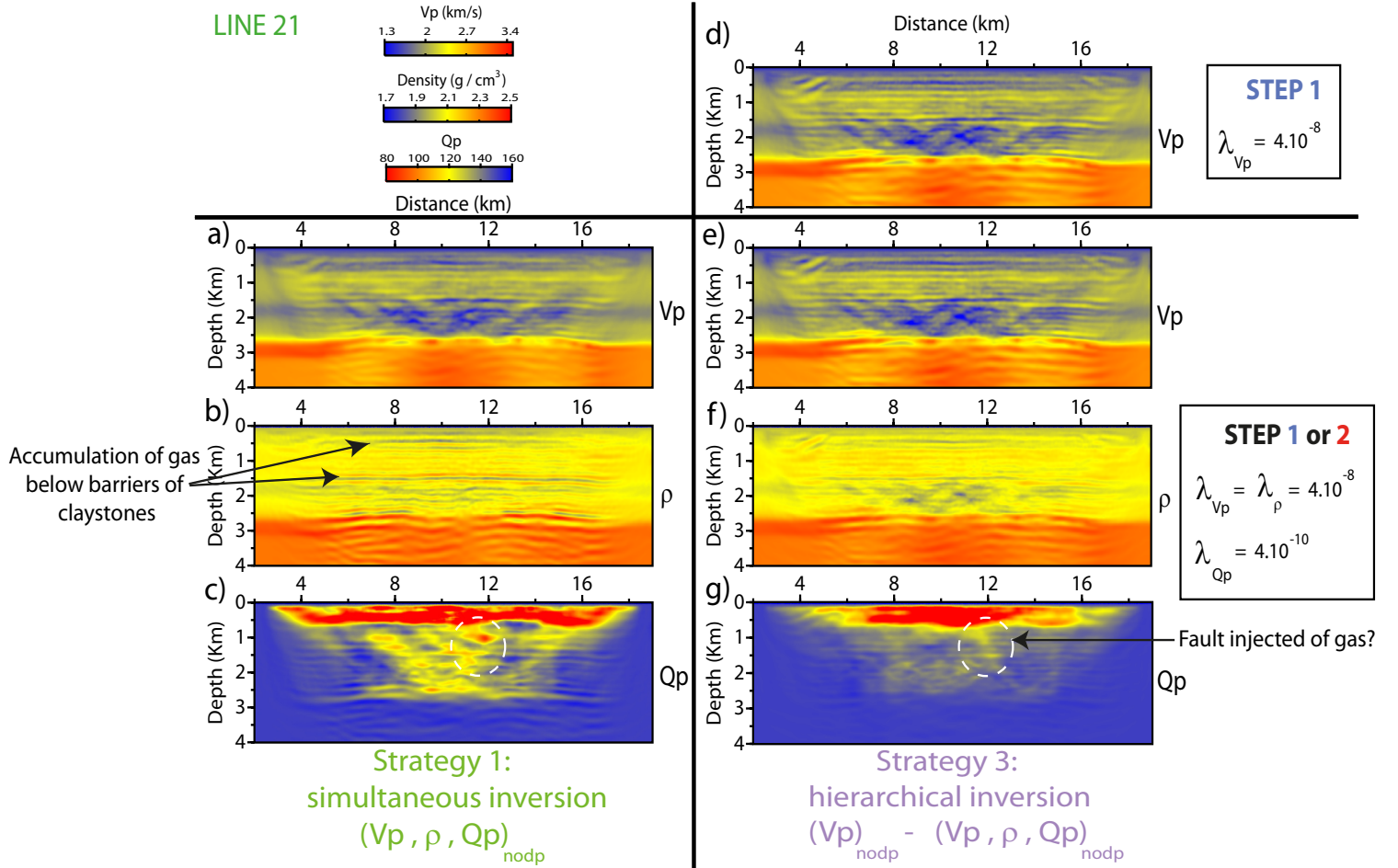


FIGURE 5.20 – FWI models for vertical velocity (a,d,e), density (b,f) and Q_P -factor (c,g) for line 21, obtained by simultaneous (a-c) or hierarchical (d-g) update of the parameters.

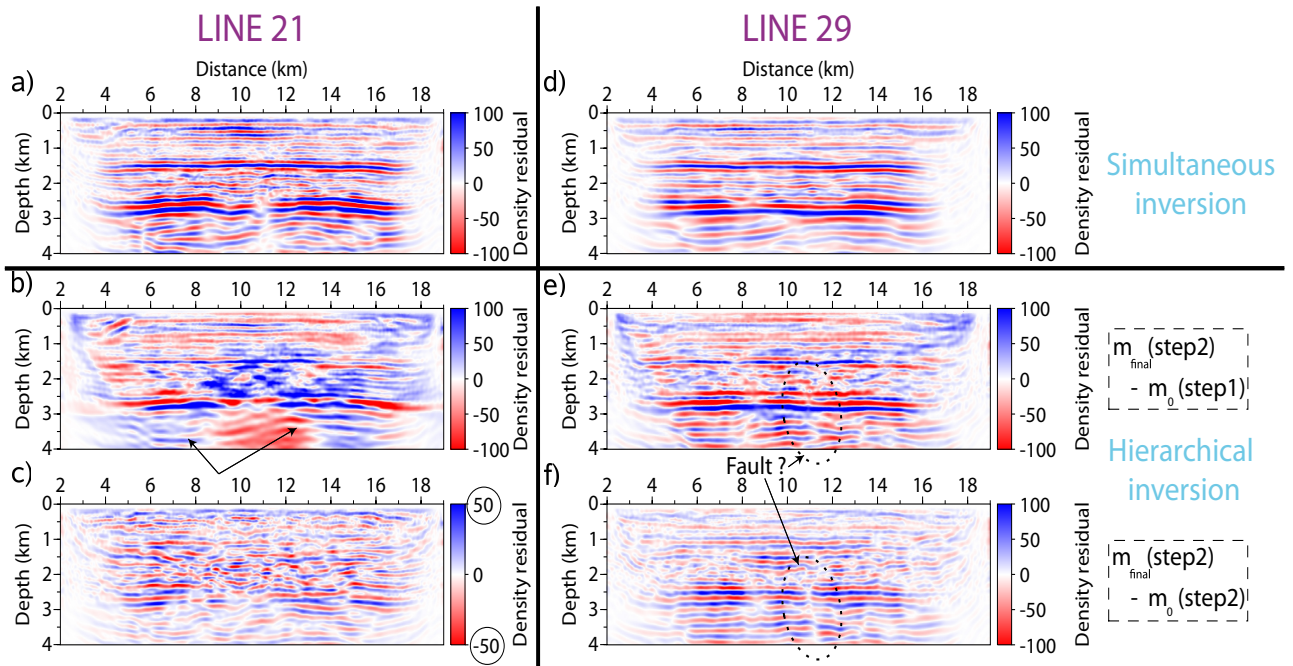


FIGURE 5.21 – Residual models of density for line 21 (a-c) and line 29 (d-f) computed by the difference between the final model and the starting model of the simultaneous inversion (a,d). For the hierarchical inversion, we distinguish the residuals derived subtracting the final model of the second step to the starting model of the first step (b,e), and the residuals representative of the contribution of the second step only (c,f). Note we changed the color scale in (c). The black arrows denote the deep reflector at 3.5 km in depth.

INVERSION VISCO-ACOUSTIQUE DES FORMES D'ONDES - RECONSTRUCTION DE LA VITESSE P, DE LA DENSITÉ OU L'IMPÉDANCE, ET DE L'ATTÉNUATION

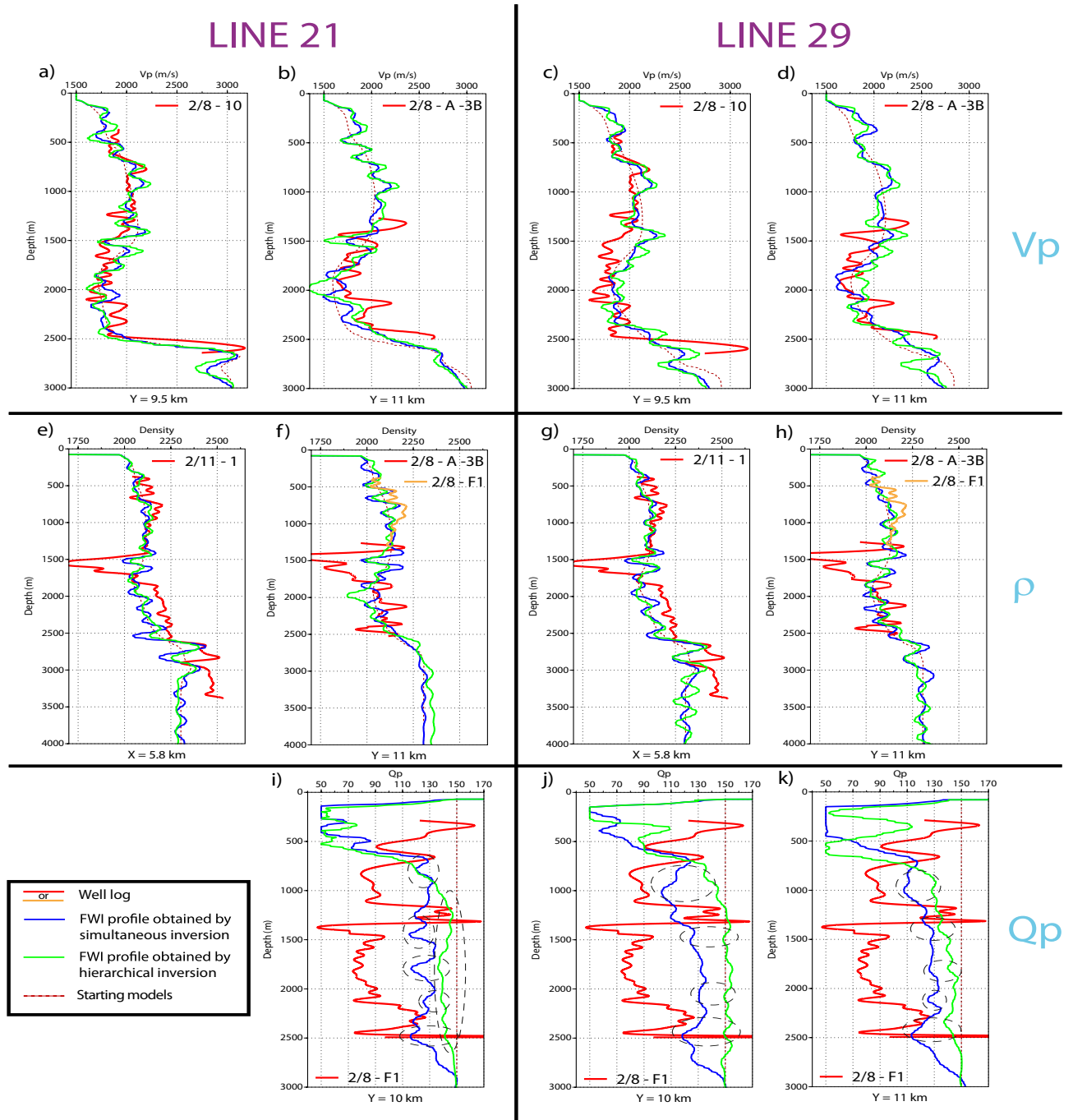


FIGURE 5.22 – Comparison of the FWI profiles of the velocity (a-d), the density (e-h) and the Q_P -factor (i-k) for line 21 (left panel) and line 29 (right panel), at the locations of the corresponding well logs (in red and orange) (see Figure 5.21). The FWI profiles are extracted from the models obtained by simultaneous inversion (in blue) or by hierarchical inversion (in green). Note the well log of Q_P is derived from the gamma-ray log using the law found in Klimentos & McCann (1990). Its magnitude is not representative of the true Q_P -factor as we arbitrarily multiplied it by 10 for an easier comparison.

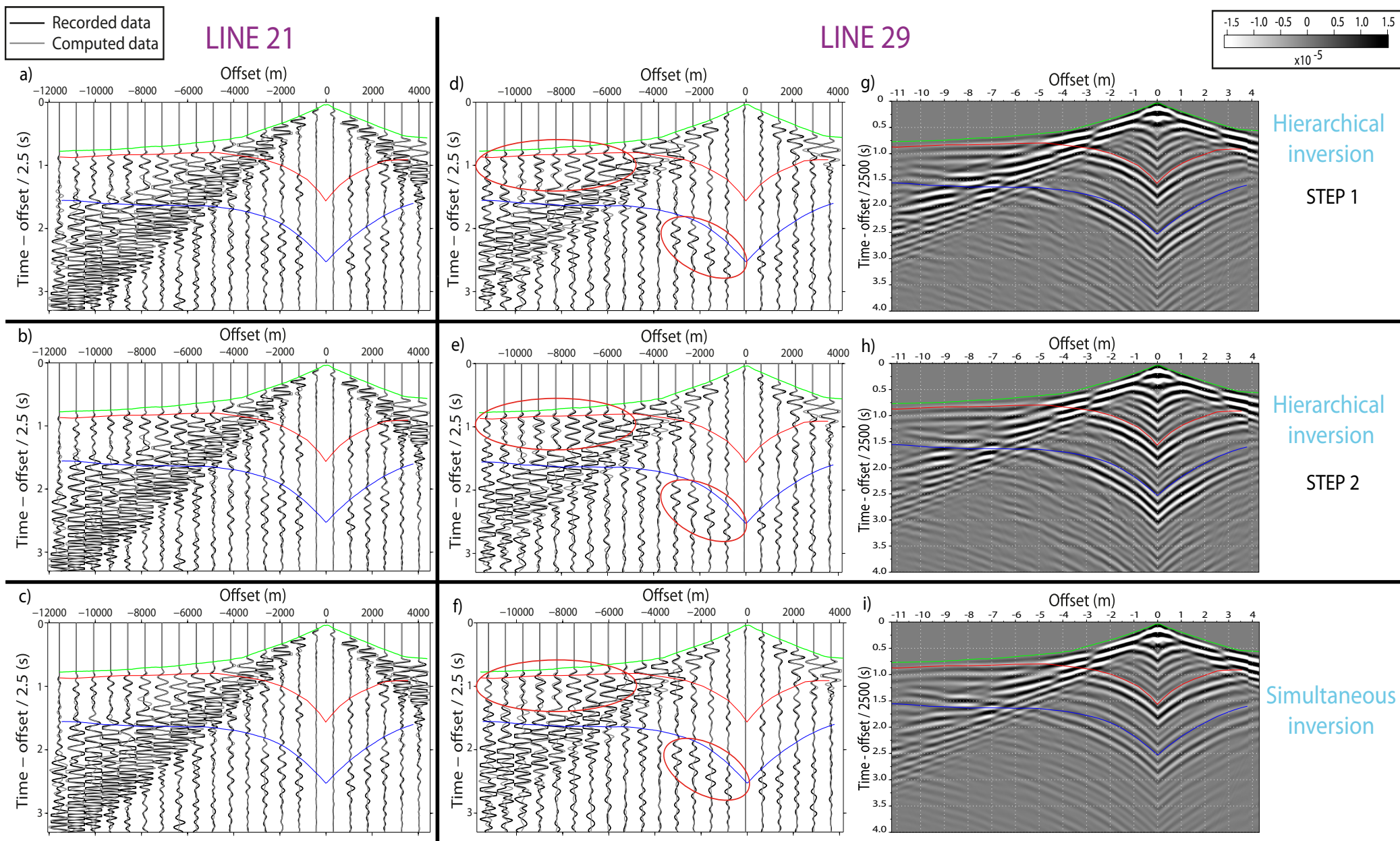


FIGURE 5.23 – (a-f) True amplitude wiggle comparison between the recorded (black line) and the computed (gray line) seismograms in the final models of the simultaneous inversion at its first step (a,d), and its second step (b,e), as well as in the final models of the simultaneous inversion (c,f), for line 21 (a-c) and line 29 (d-f). (g,i) Same computed data than in (d,f) but displayed with a black/white scale. The green, red and blue curves denotes the manually picked first arrival traveltimes and the reflections from the top and the bottom of the gas.

5.7 Conclusion

We have presented an application of visco-acoustic FWI of hydrophone data for imaging the compressional wavespeed, the density and the Q_P -factor in marine environments. A first recommendation is to combine the density with the compressional wavespeed in the subsurface parametrization rather than the impedance. The main reason is that in the first case, the radiation pattern of the compressional wavespeed is isotropic leading to a high resolution reconstruction of the velocity, which is the parameter with the dominant imprint in the data. A second recommendation is to update the velocity and the density in a hierarchical manner rather than simultaneously. This hierarchical approach contributes to update the density during the second step of the inversion in a more stable manner than the simultaneous update of the velocity and the density in one go. Imaging the attenuation in marine environment leads to additional difficulties associated with the small footprint of this parameter in the data. We found that a judicious tuning of the FWI regularization was a key issue to scale properly the model perturbations of the velocity, density and Q_P -factor. Here, by judicious tuning is meant that the damping terms in the Hessian should be scaled to the diagonal terms of the Hessian matrix to take into account the variable sensitivity of the data to the variable parameter classes. In the case of the imaging of velocity, density and Q_P -factor, we recommend also a hierarchical approach where the dominant wave speed parameter is reconstructed first before the simultaneous update of the three parameter classes. The update of the P-wave velocity, density and Q_P -factor by visco-acoustic FWI from the hydrophone data recorded at the Valhall field can provide reliable starting and background models for subsequent elastic FWI for imaging the compressional and shear wavespeeds, as shown in the next publication.

Acknowledgments

This study was funded by the SEISCOPE consortium <http://seiscope.oce.eu>, sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The linear systems were solved with the MUMPS package, available on <http://graal.ens-lyon.fr/MUMPS/index.html> and <http://mumps.enseiht.fr>. The mesh generation was performed with help of TRIANGLE, available on <http://www.cs.cmu.edu/~quake/triangle.html>. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and of CINES/IDRIS under the allocation 2010- [project gao2280] made by GENCI (Grand Equipement National de Calcul Intensif), and we gratefully acknowledge both of these facilities and the support of their staff. We also thank BP Norge AS and their Valhall partner Hess Norge AS for allowing access to data to the Valhall data set as well as the well-log velocities.

5.8 Conclusion à l'article

Nous avons décrit dans ce chapitre l'origine des difficultés et les pistes connues pour mieux contraindre la reconstruction de multiples classes de paramètres, que nous avons illustré à travers des applications au cas d'étude de Valhall. La comparaison des paramétrisations (V_P, ρ) et (V_P, I_P) a montré l'importance de considérer une paramétrisation conférant au paramètre

dominant un diagramme de rayonnement "large bande" (ou isotrope), et a permis d'illustrer les problèmes liés aux différences de dynamiques d'échelle. La première paramétrisation en vitesse et densité est celle qui donne les meilleures reconstructions à la fois pour la vitesse, la densité et l'impédance. Cependant, une perspective pour la seconde paramétrisation serait de rajouter un terme de contrainte dans la fonction coût contraignant la stabilité de la densité déduite des deux paramètres inversés, en bornant par exemple les valeurs prises par la densité, ou encore de façon à ce que les modèles de densité soient suffisamment lisses. La comparaison entre les stratégies d'inversion simultanées et hiérarchiques tend à montrer qu'il est préférable d'inverser la vitesse avant les paramètres secondaires de densité et d'atténuation. L'inversion simultanée a tendance à surestimer les paramètres, alors que l'inversion hiérarchique les sous-estime au bénéfice d'un gain en résolution, permis par la création préalable d'un modèle de vitesse expliquant la cinématique des différentes phases. La corroboration des résultats obtenus avec les deux méthodes permet de se faire une idée plus précise des structures réelles ou artificielles observées. Nous avons aussi illustré la faible empreinte qu'ont la densité et l'atténuation dans les données, ainsi que le faible impact qu'ont les modèles initiaux de ces paramètres secondaires sur la qualité de reconstruction de la vitesse seule. Par ailleurs, dans le cadre d'une régularisation Tikhonov, nous recommandons l'utilisation d'un amortissement d'inversion adapté à la sensibilité de chaque paramètre. La gamme de valeurs d'amortissement possible est trouvée par essai erreur en se basant sur notre a priori géologique. L'application de notre protocole d'inversion multiparamètres acoustique sur le champ de Valhall a abouti à la création de modèles géologiquement plausibles pour lesquels des interprétations peuvent être proposées. L'évaluation de la qualité des résultats est réalisée en regardant l'accord aux données, et le caractère vraisemblable des modèles est mis à l'épreuve en les comparant avec des logs de puits. Lorsque de nombreux modèles d'inversion sont obtenus à partir de différentes stratégies d'inversion, de paramétrisations, de choix des amortissements, leur moyenne pondérée par des informations a priori pourrait se révéler primordiales pour mieux contraindre les solutions les plus probables parmi la multitude possible dans l'espace des modèles.

Chapitre 6

Inversion visco-élastique des formes d'ondes - Reconstruction des vitesses P et S

Sommaire

6.1	Introduction à l'article	222
6.2	Résumé	222
6.3	Introduction	223
6.4	Inversion des formes d'ondes visco-élastique de données multi-composantes	225
6.5	Stratégies multi-échelles et paramétrisation de l'inversion des formes d'ondes	228
6.5.1	Approches hiérarchiques pour une inversion multiparamètre	228
6.5.2	Choix de la paramétrisation et ses motivations	228
6.5.3	Diagrammes de rayonnement des paramètres V_P et V_S	229
6.5.4	Modèles initiaux et stratégies d'inversion dans cette étude	231
6.6	Protocole pour les données et l'inversion des formes d'ondes . . .	232
6.7	Résultats d'inversion	234
6.8	Evaluation de la qualité des modèles	234
6.8.1	Comparaison avec des logs soniques	234
6.8.2	Accord avec les données	236
6.8.3	Estimation des ondelettes source	239
6.9	Dérivation d'autres quantités physiques et interprétations géologiques	248
6.9.1	Structure superficielle du modèle de vitesses d'ondes S	248
6.9.2	Coefficient de Poisson et autres combinaisons de V_P et V_S	248
6.9.3	Paramètres I_P, I_S, Q_S	253
6.9.4	Indicateurs d'hydrocarbures	253
6.10	Conclusion	255
6.11	Conclusion aux chapitres multiparamètres	256

6.1 Introduction à l'article

Je présente dans ce chapitre une inversion des formes d'ondes pour l'obtention des paramètres V_{P0} et V_S sur le champ de Valhall, appliquée aux jeux de données 4C des lignes 21 et 29. Les modèles initiaux de V_{P0} , de densité et d'atténuation utilisés pour l'inversion élastique ont été obtenus dans le chapitre précédent par inversion des formes d'ondes acoustique anisotrope. Les paramètres de Thomsen δ et ϵ sont construits par tomographie en réflexion et ne font pas partie des paramètres à inverser. J'applique une stratégie hiérarchique consistant à inverser la composante d'hydrophone, en vue d'utiliser l'information AVO contenue dans les réflexions PP, suivie de l'inversion des composantes de vitesses du géophone pour exploiter les converties PS. Je montre l'intérêt de cette approche grâce à différents critères de qualité, relativement au test où sont directement inversées les données de géophone. La qualité des résultats est évaluée grâce à la comparaison de V_S avec des logs de puits, grâce au calcul de sismogrammes dans les modèles obtenus et à l'interprétation des arrivées converties dans les données, ainsi que par l'estimation d'ondelettes sources par tir pour chaque composante. La stratégie hiérarchique donne de meilleurs résultats en raison des nombres d'ondes intermédiaires fournis à l'imagerie lors de la première étape, que l'inversion des converties PS seule ne peut fournir en raison des faibles vitesses S du milieu. Les paramètres obtenus par FWI nous ont servi dans une dernière partie à dériver via des relations empiriques diverses quantités pétrophysiques du milieu, nous permettant de proposer des interprétations géologiques à nos résultats.

**Multiparameter full waveform inversion of multicomponent OBC
data from the Valhall field.**

Part 2 : imaging compressive and shear-wave velocities.

V. Prieux, R. Brossier, S. Operto, J. Virieux
For submission to Geophysics

6.2 Summary

In this study, we assess the feasibility of 2D vertical transverse isotropic (VTI) visco-elastic full waveform inversion (FWI) for the reconstruction of highly resolved models of the vertical P-wave and S-wave velocities (V_{P0} and V_S) from 4C OBC wide-aperture data recorded in the Valhall oil field. Modeling is carried out by a frequency domain Discontinuous Galerkin method, and we proceed the inversion following a multiscale approach from the low to high frequencies. For performing reliable multi-parameters inversion of the multi-components data, we propose a visco-acoustic & visco-elastic FWI workflow, which proceeds hierarchically over data components and parameter classes. First, our strategy relies on the reconstruction of the dominant parameters by visco-acoustic FWI. We show in the companion article the reconstruction of V_{P0} followed by the density (ρ) and the quality factor (Q_P). We now consider the updating of the second-order parameter (V_S) by visco-elastic FWI. Second, data components (hydrophone versus geophone, wide apertures versus short apertures) are progressively introduced in the elastic inversion, such that the long wavelengths of V_S are reconstructed before the shorter ones, hence, introducing a second level of data preconditioning in our multiscale approach. The FWI models are validated against sonic logs, seismic modeling, and source wavelet estimation. Petro-physical quantities inferred from the V_{p0} , density and V_s FWI models highlight the re-

servoir location, the main discontinuities of the subsurface below which accumulates gas, and suggests the presence of gas injected fractures.

6.3 Introduction

Subsurface lithology and reservoir characterization require quantitative estimation of physical rock properties. The joint knowledge of the P and S-wave velocities is of primary interest for better fluid specification (Gregory, 1976), as a potential indicator of hydrocarbon reservoir (Tatham & Stoffa, 1976), or to estimate porosity ϕ and grain size through empirical relations of the form $1/V = A + B\phi$, where V denotes the compressional or shear wavespeed (Domenico, 1984). Shear-wave velocity models can also be used for the processing of multicomponent seismic data, as for wavefield decomposition or static corrections (Muyzert, 2000). Full waveform inversion (FWI) (Virieux & Operto, 2009) is nowadays often used to derive a P-wave velocity background model in the acoustic approximation for reverse-time-migration (Sirgue *et al.*, 2010). Yet, multi-parameter FWI based on the visco-elastic wave equation provides the necessary framework to evolve beyond the velocity model building task and derive more realistic multi-parameter subsurface models, as promoted by Tarantola (1986); Brossier (2011a). Application of elastic FWI to real data remains challenging, because of the computational cost of the elastic modeling that is one order of magnitude more expensive than its acoustic counterpart, and the strong non-linearities of the elastic multi-parameter inversion. Few applications of elastic FWI to marine and land dataset were presented in the past, where the compressional and/or the shear wavespeeds, the elastic impedance or the Poisson's ratio are reconstructed from short-spread reflection data (less than 5 km offset) (Crace *et al.*, 1990, 1992; Sun & McMechan, 1992; Igel *et al.*, 1996; Djikpéssé & Tarantola, 1999a). Application of FWI to long-spread streamer data was presented by Freudenreich *et al.* (2001) and Shipp & Singh (2002), although only the P wavespeed was updated during inversion in the last study.

In this study, we present an application of multi-parameter elastic FWI to multi-component wide-aperture data recorded by ocean bottom cables in the Valhall field. The reconstruction of a first compressional velocity model, a density model and an attenuation model from the hydrophone component in the acoustic approximation is presented in a companion study referred to as Paper I in the following.

Multi-component wide-aperture acquisition geometries is a key for the success of multiparameter elastic FWI in addition to the recording of low frequencies (Manukyan *et al.*, 2012). It is well acknowledged that short-spread streamer acquisitions do not allow for the reconstruction of the intermediate wavelengths of the subsurface because of insufficient wide-angle illumination (Jannane *et al.*, 1989; Neves & Singh, 1996). This justifies the scale separation underlying the two main tasks of conventional seismic reflection processing workflow : velocity macro-model building and imaging of the reflectivity by migration. Wide-aperture geometries provide a suitable framework to build the long and intermediate wavelengths of the compressional and shear wavespeeds from long-offset data and hence continuously sample the wavenumber spectrum of the subsurface. Second, wide-aperture multi-component recordings are necessary to record all of the scattering modes (P-P, P-S, S-P and S-S diffractions) over a sufficiently-broad range of diffraction angles, that should help to avoid the trade-off between the update of the compressional and shear wavespeeds. This is even more critical in marine environments where the waves should undergo at least one P-to-S conversion to allow for S-wave recording

by sea-bottom acquisitions. The amount of P-S conversion can be furthermore quite limited in the case of soft-seabed environment. Another key issue related to S-wave velocity imaging is the building of a sufficiently-accurate starting model for full waveform inversion. One difficulty results from the smaller wavelengths of the shear wavefields, which make the reconstruction of the shear velocity subject to cycle skipping artifacts and can prevent the continuous sampling of the wavenumber spectrum of the shear velocity model. A second difficulty is related to the picking of reliable P-S converted phases before reflection traveltime tomography.

A feasibility analysis of elastic FWI of multi-component ocean-bottom seismic data was presented by Sears *et al.* (2008). They propose a hierarchical strategy where the long and intermediate wavelengths of the shear wavespeeds are first updated from the information content of the hydrophone component associated with the AVO of the P-P reflections. Sensitivity of the PP reflections to contrasts in the shear impedance (I_S) is one order of magnitude smaller than to contrasts in the compressive impedance (I_P) (Igel & Schoenberg, 1995). This strategy is however of particular interest for marine case. The first reason being that soft-seabed marine environments generally prevent significant PS conversions at the sea bed, which are useful indirect measurements of the shear wavespeed, and secondly because marine medium are characterized by (very) slow shear velocity models preventing the reconstruction of the large and intermediate spatial wavenumbers. Building a reliable P-wave velocity model during an earlier stage consisting of traveltime tomography or acoustic FWI can help to mitigate the risk of cycle-skipping artifacts during the inversion of the P-P reflections for the updating of the shear velocities. In a second step, the shorter wavelengths of the shear velocity models can be updated from the P-S and S-S scattering modes recorded by the geophone components. This approach has been applied to real data in Sears *et al.* (2010). In this study, we shall follow a similar approach than the one promoted by Sears *et al.* (2008).

Before showing the application of elastic FWI to the real data from the Valhall field, it is worth reviewing the main results of the FWI of elastic data computed in the synthetic Valhall model (Brossier *et al.*, 2009c). Comparison between seismograms computed in the acoustic and elastic approximations has shown that the seismic wavefield is dominated by the compressional wavefield. In this framework, a first acoustic inversion of the elastic data recorded by the hydrophone component provides a reliable P-wave velocity model. This is consistent with the conclusions of Barnes & Charara (2009), who have shown the applicability of acoustic FWI in soft seabed environments. In a second step, the P-wave and the S-wave velocity models can be updated jointly from the geophone components. V_P was well recovered from the first step, but V_S suffered (especially in depth) of a lack of low wavenumbers, attributed to its low velocities (Figure 6.1(d-f)). It has been verified that the horizontal velocity component was mandatory to obtain a significant update of the shear-wave velocities. In this study, we shall see that the additional intermediate stage promoted by Sears *et al.* (2008), which consists of updating the P-wave and the S-wave velocities from the hydrophone data will contribute to improve yet the results.

In the first part of this study, we review the main features of the elastic FWI method that we use. Then, we discuss several hierarchical strategies which can be viewed to perform multiparameter FWI of multicomponent data. In particular, the choice of the subsurface parametrization for FWI, the hierarchy with which the different classes of parameters are updated and the hierarchy with which the different data components are inverted are three key issues, which are addressed. After the presentation of the data and the FWI setup, we present the FWI models and we discuss their relevance based on validation against sonic logs, seismic mo-

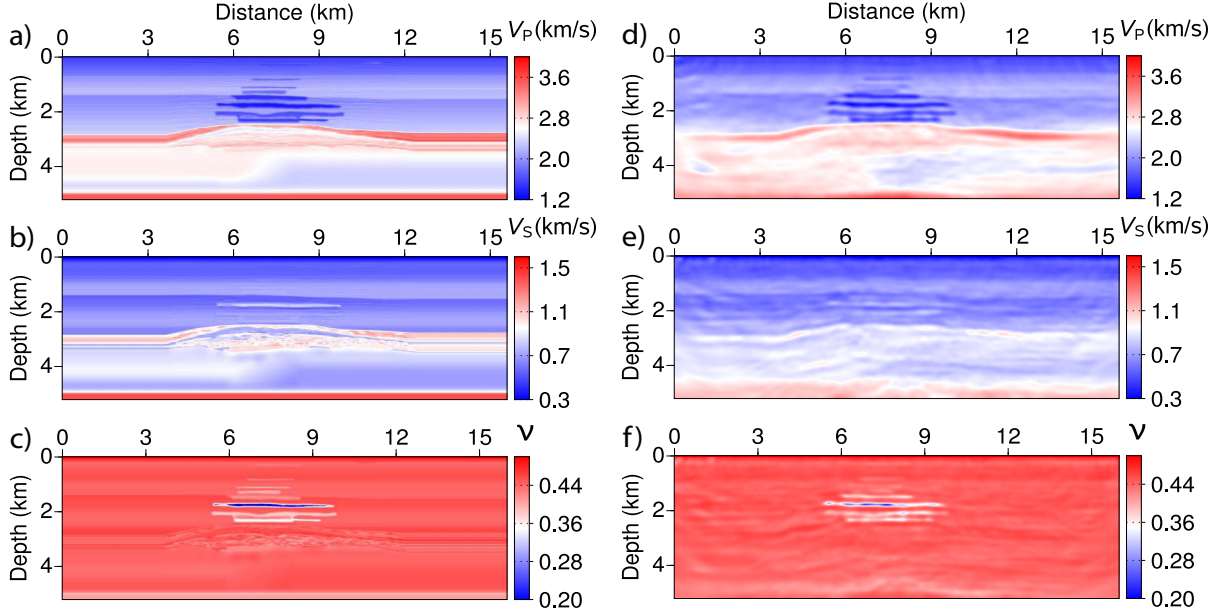


FIGURE 6.1 – (a-c) True V_P (a), V_S (b) and Poisson's ratio (c) from the synthetic Valhall case. (d-f) Final results inferred from the hierarchical strategy from Brossier *et al.* (2009c). Final V_P (d), V_S (e) FWI models, and the corresponding Poisson's ratio (f).

deling and source wavelet estimation. Finally, we derive some petro-physical quantities from the elastic FWI models, that are used to interpret several geological features.

6.4 Visco-elastic full-waveform inversion of multi-component data

In this study, we seek to reconstruct the vertical compressional wavespeed and the shear wavespeed (denoted by V_{P0} and V_S , respectively) by visco-elastic anisotropic FWI from 2D multi-component sensors composed of one hydrophone and two horizontal and vertical geophones. The subsurface is assumed to be a visco-elastic vertically transverse isotropic (VTI) medium.

Seismic modeling is performed in unstructured triangular meshes with a frequency-domain velocity-stress finite-element discontinuous Galerkin method (Brossier *et al.*, 2010b).

Full waveform inversion (FWI) is performed in the frequency domain by local optimization, where the gradient of the misfit function is computed with the adjoint-state method (Plessix, 2006; Chavent, 2009). A forward-modeling operator based on the second-order wave equation for particle velocities and piecewise constant interpolation function (P0) in the finite-element method is used to build the diffraction kernel of the FWI from self-adjoint operators, while the incident and the adjoint wavefields are accurately computed with a first-order velocity-stress formulation of the wave equation with either P0, P1 (linear) or P2 (quadratic) interpolation functions (Brossier, 2011a).

The misfit function $\mathcal{C}(\mathbf{m})$ is given by

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \sum_{j=1}^{N_c} \Delta \mathbf{d}_j^\dagger \mathbf{W}_{d_j} \Delta \mathbf{d}_j + \frac{1}{2} \sum_{i=1}^{N_p} \lambda_i (\mathbf{m}_i - \mathbf{m}_{prior_i})^\dagger \mathbf{W}_{m_i} (\mathbf{m}_i - \mathbf{m}_{prior_i}), \quad (6.1)$$

where $\Delta \mathbf{d} = (\Delta \mathbf{d}_1, \dots, \Delta \mathbf{d}_{N_c})$ and N_c denote the multi-component data-residual vector and the number of data components involved in the inversion, respectively. In this study, N_c equals to 1 or 2 depending whether the hydrophone or the two geophone components are considered during the inversion. Each component j of the data residual vector is weighted by the weighting matrix \mathbf{W}_{d_j} . This matrix can weight data residuals according to the standard error and/or according to source-receiver offset (Ravaut *et al.*, 2004; Operto *et al.*, 2006). The multiparameter subsurface model is denoted by $\mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_{N_p})$, where N_p denotes the number of parameter classes. In this study, the two classes of parameter are the compressional vertical wavespeed and the shear wavespeed, which can be updated jointly or hierarchically as we shall see. A Tikhonov regularization is applied to each class of parameter i through a roughness operator \mathbf{W}_{m_i} , which forces the difference between the model \mathbf{m}_i and a prior model \mathbf{m}_{prior_i} to be smooth. The scalars λ_i control the weight of the data-space misfit function $\frac{1}{2} \sum_{j=1}^{N_c} \Delta \mathbf{d}_j^\dagger \mathbf{W}_{d_j} \Delta \mathbf{d}_j$ relatively to the model-space misfit functions $(\mathbf{m}_i - \mathbf{m}_{prior_i})^\dagger \mathbf{W}_{m_i} (\mathbf{m}_i - \mathbf{m}_{prior_i})$. Of note, the scalar λ_i can be adapted to each parameter class i , that is helpful to account for the variable sensitivity of the data to each parameter class as shown in the Paper I.

The solution for the perturbation model, that minimizes the misfit function at iteration k is given by

$$\Delta \mathbf{m}^{(k)} = \Re \left(\widehat{\mathbf{W}}_m^{-1} \mathbf{J}^{(k)t} \mathbf{W}_d \mathbf{J}^{(k)*} + \widehat{\mathbf{W}}_m^{-1} \left(\frac{\partial \mathbf{J}^{(k)t}}{\partial \mathbf{m}^t} \right) (\Delta \mathbf{d}^{(k)*} \dots \Delta \mathbf{d}^{(k)*}) + \Lambda \right)^{-1} \Re \left(\widehat{\mathbf{W}}_m^{-1} \mathbf{J}^{(k)t} \mathbf{W}_d \Delta \mathbf{d}^{(k)*} + \Lambda (\mathbf{m}^{(k)} - \mathbf{m}_{prior}) \right), \quad (6.2)$$

where Λ is a block diagonal damping matrix :

$$\Lambda = \begin{pmatrix} \lambda_1 \mathbf{I}_M & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & \lambda_{N_p} \mathbf{I}_M \end{pmatrix} \quad (6.3)$$

where \mathbf{I}_M is the identity matrix of dimension M , the number of nodes in the computational mesh. In equation 6.2, \Re is the real part of a complex number. The matrices $\widehat{\mathbf{W}}_d$ and $\widehat{\mathbf{W}}_m$ are $N_c \times N_c$ and $N_p \times N_p$ block diagonal matrices, where each block is formed by the \mathbf{W}_{d_j} and \mathbf{W}_{m_i} matrices, respectively. The term, whose inverse must be computed in equation 6.2 is the full Hessian, which contains 3 terms. The first one aims to correct the gradient for linear effects such as the limited bandwidth of the source, the limited spread of the acquisition geometry, and the geometrical spreading of the data. The second term corrects the Hessian from double-scattering effects in nonlinear inverse problems (Pratt *et al.*, 1998). The third regularization term damps the deconvolution action of the first two terms of the Hessian to improve the conditioning of the Hessian matrix. The term on the right-hand side of the inverse of the Hessian in equation 6.2 is the gradient of the misfit function, which is composed of two terms : the first one represents the contribution of the data and is formed by the zero-lag correlations

of the partial derivative wavefields at the receiver positions with the data residuals, while the second term is the model-space regularization term.

It is worth reminding that the partial derivative wavefields satisfy

$$\mathbf{B}(\omega, \mathbf{m}(\mathbf{x})) \frac{\partial \mathbf{v}(\omega, \mathbf{x})}{\partial m_i} = - \frac{\partial \mathbf{B}(\omega, \mathbf{m}(\mathbf{x}))}{\partial m_i} \mathbf{v}(\omega, \mathbf{x}), \quad (6.4)$$

where $\mathbf{B}(\omega, \mathbf{m}(\mathbf{x}))$ is the forward-problem operator (in our case, the second-order wave equation operator for particle velocities) and \mathbf{v} is the incident particle velocity wavefield. The right-hand side of equation 6.4 is the secondary virtual source of the partial derivative wavefield, the spatial support and the temporal support of which are centered on the position of the diffractor m_i and on the arrival time of the incident wavefield at the diffractor m_i , respectively (Pratt *et al.*, 1998). The diffraction or radiation pattern of this virtual source is given by $\frac{\partial \mathbf{B}(\omega, \mathbf{m}(\mathbf{x}))}{\partial m_i}$ and gives some insights on the sensitivity of the data to the parameter m_i as a function of the diffraction (or aperture) angle.

We use the quasi-Newton L-BFGS optimization algorithm to solve equation 6.2 (Nocedal, 1980; Nocedal & Wright, 1999). The L-BFGS algorithm computes recursively an approximation of the product of the inverse of the Hessian with the gradient, from few gradients and few solution vectors from the previous iterations. As preconditioner of this iterative search, we use a diagonal approximation of the approximate Hessian (the linear term) damped by the Λ matrix,

$$PreH = \left(\widehat{\mathbf{W}}_m^{-1} \text{diag} \left\{ \mathbf{J}^{(k)t} \mathbf{W}_d \mathbf{J}^{(k)*} \right\} + \Lambda \right)^{-1}. \quad (6.5)$$

Of note, the smoothing operator $\widehat{\mathbf{W}}_m^{-1}$ is applied to the first two terms of the Hessian in equation 6.2 for implementation convenience, although another implementation consists of using $\widehat{\mathbf{W}}_m$ in the third term of the Hessian. Similar comment applies to the gradient term, where the operator $\widehat{\mathbf{W}}_m$ or its inverse can be applied to the first or second terms of the gradient accordingly. Our implementation allows the preconditioner of the L-BFGS algorithm, equation 6.5, to be diagonal and hence easy to invert because the matrix Λ is diagonal unlike $\widehat{\mathbf{W}}_m$.

The smoothing operators $\mathbf{W}_{m_i}^{-1}$ are exponential function given by :

$$\mathbf{W}_{m_i}^{-1}(z, x, z', x') = \sigma_i^2(z, x) \exp\left(\frac{-|x - x'|}{\tau_x}\right) \exp\left(\frac{-|z - z'|}{\tau_z}\right), \quad (6.6)$$

where τ_x and τ_z denote the horizontal and vertical correlation lengths, defined as a fraction of the local wavelength. The coefficient σ_i represents the standard error and is scaled to the physical unit of the parameter class i . This exponential function is used for $\mathbf{W}_{m_i}^{-1}$ because its inverse in the expression of the misfit function can be computed analytically (Tarantola (1987), section 7.21, pages 308-309).

6.5 Multiscale strategies and FWI parametrization

6.5.1 Hierarchical approaches for multiparameter inversion

Many hierarchical FWI strategies can be viewed to reduce the nonlinearity of the inversion when multi-component data are inverted to reconstruct multiple classes of parameters (e.g. Tarantola, 1986). A data-driven level of hierarchy can be designed by proceeding from the inversion to the low frequencies to the higher ones such that the large wavelengths of the subsurface are reconstructed before the smaller ones in the framework of multiscale imaging (e.g. Bunks *et al.*, 1995; Sirgue & Pratt, 2004). Additional data preconditioning can be viewed to inject more resolving or more complex information in the inversion by proceeding from early arrivals to later ones (Brossier *et al.*, 2009b) or by proceeding from short offsets to longer ones in the framework of layer-stripping strategies (Shipp & Singh, 2002; Wang & Rao, 2009). In the framework of multicomponent data, the hierarchy with which data components are injected in the inversion is also a key feature since the sensitivity of the data to one model parameter can depend on the data component (hydrophone versus geophone) (Sears *et al.*, 2008, 2010). For example, the sensitivity of the horizontal-geophone data to the shear wavespeed is higher than the one of the vertical-geophone or hydrophone data.

When multiparameter inversion is performed, additional levels of hierarchy in the model space can also be considered during FWI. First, one may choose to image those parameters, which have a dominant imprint in the data, before imaging secondary parameters (Tarantola, 1986). Second, the model parameters, which are sensitive to the wide-aperture data, can be reconstructed before the ones, which are sensitive to the short-aperture data. According to the relationship between the aperture angles and the wavenumber coverage of the model (Miller *et al.*, 1987; Jin *et al.*, 1992; Pratt & Worthington, 1990), the long wavelengths of the first category of parameters should be reconstructed before the short wavelengths of the second category of parameters, hence honoring the multiscale approach of FWI. An illustrative example of such parametrization involves the wavespeed and the impedance, which scatter waves at wide and short apertures, respectively, when these two classes of parameters are combined in the FWI parametrization.

6.5.2 Choice of parametrization and its motivations

The choice of the subsurface parametrization should be driven by at least three criteria : the sensitivity of the data to the parameter with respect to the aperture angle should be as broad as possible to guarantee a high-resolution reconstruction of the parameter. On the other hand, the trade-off between parameter classes should be reduced as much as possible, that requires the radiation patterns of each parameter class to not overlap significantly as a function of the aperture angle. A suitable trade-off between these two criteria should be clearly found. Third, the sensitivity of the data to the parameter for a significant range of aperture angles must be sufficiently high such that the information content of the data can be extracted from noise during inversion. In this study, the subsurface model is parametrized by the vertical compressional wavespeed, the shear wavespeed, the density, the quality factors Q_P and Q_S and the Thomsen parameters. Other possibilities involve the P and S slowness, the P and S-wave impedances, the Poisson coefficient or the elastic moduli (Forgues & Lambaré, 1997).

Test	Step	Inverted data	Inverted parameter	Approximation
	1	hydrophone	V_{P0}	VTI visco-acoustic
	2		V_{P0}, ρ, Q_P	
G	3	geophone	V_{P0}, V_S	VTI visco-elastic
H	3	hydrophone		
HG	4	geophone		

TABLE 6.1 – Name given to our different tests. Test G denotes the inversion when are directly inverted the geophone data, while Test H and HG determine the two steps of the hierarchical strategy, when are inverted the hydrophone followed by the geophone.

A hierarchical reconstruction of the vertical compressional wavespeed, the density and attenuation has been performed in the visco-acoustic approximation for computational savings and is described in Paper I. We have followed a hierarchical approach where a model of the compressional wavespeed is reconstructed during a first step before the joint update of this wavespeed, the density and the attenuation during a second step (Table 6.1). Of note, we did not use the parametrization promoted by Tarantola (1986), who recommends to reconstruct first the long wavelengths of the P wavespeed from the wide apertures before reconstructing the short wavelengths of the P impedance from the short apertures. We rather favor a parametrization for which the dominant parameter (namely, the P wavespeed) has an isotropic radiation pattern, because we found that the high-resolution model of the wavespeed obtained close of a first inversion step was useful for the reconstruction of the density and attenuation during a second step. Although elastic inversion is generally preferable, the footprint of the shear wave velocities is sufficiently small in the hydrophone data to build a reliable V_P velocity model in the visco-acoustic approximation (Brossier *et al.*, 2010c; Prieux *et al.*, 2011). The synthetic tests performed by Igel *et al.* (1996) in the (V_P, I_P, ν) parametrization have shown the Poisson's ratio inferred from the P wave AVO is very sensitive to the inaccuracies of I_P , even when V_P is known. Debski & Tarantola (1995) states the choice of the (V_P, V_S, ρ) parametrization is a mistake to invert V_S from this AVO information, because it leads to strongly correlated uncertainties. He advocates to use instead other parametrizations like (I_P, ν, ρ) or (I_P, I_S, ρ) for which the posterior probability density is maximally decoupled between the three parameters, with the largest error left on the density parameter. However, we argue V_P and ρ were rather satisfyingly determined by a previous acoustic FWI, and that the density is kept fixed during the elastic inversion.

6.5.3 V_P and V_S radiation patterns

The radiation patterns of the compressional and shear wavespeeds, which are computed in the asymptotic ray+Born approximation for the four diffraction modes P-P, P-SV, SV-P and SV-SV, are shown in Figure 6.2 when an elastic isotropic medium parametrized by V_P , V_S and the density is considered (Forgues & Lambaré, 1997). A V_P model perturbation generates only P-P diffraction with an isotropic radiation pattern (Figure 6.2a). Therefore, a high-resolution reconstruction of the dominant parameter V_P is expected with the parametrization (V_P, V_S, ρ) . A V_S model perturbation generates P-P diffraction at intermediate aperture angles with smaller amplitudes than the one generated by a V_P model perturbation. Therefore, a smooth recons-

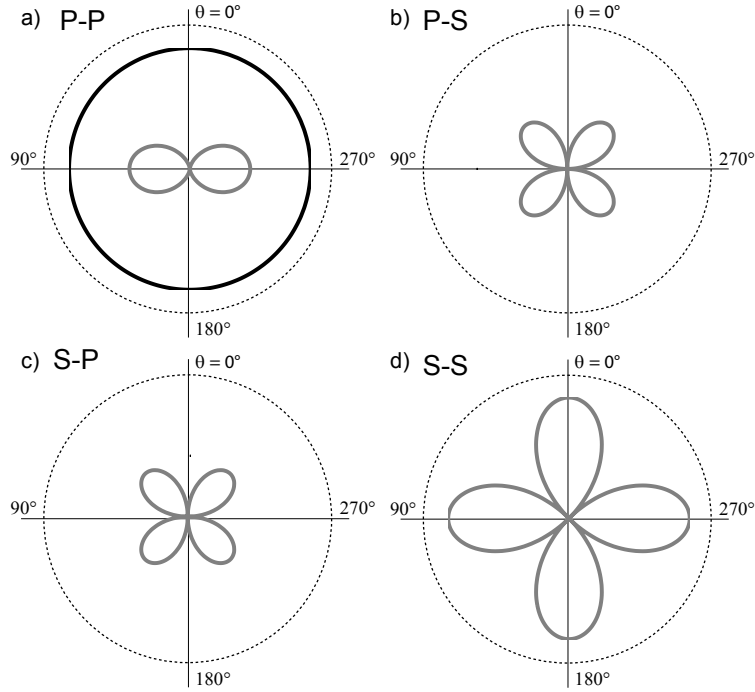


FIGURE 6.2 – Radiation pattern of compressional-wave (black) and shear-wave (gray) velocity parameters for the P-P (a), P-SV (b), SV-P (c) and SV-SV (d) modes computed in the ray+Born approximation (Forgues & Lambaré, 1997). The radiation patterns are plotted in polar coordinates as a function of the scattering (or aperture) angle. Note that a P-wave velocity perturbation generates only P-P diffraction.

truction of the V_S parameter is expected from the hydrophone component, which is consistent with the results of Sears *et al.* (2008, 2010). The sensitivity of the hydrophone data to the V_S parameter (Aki & Richards, 1980) is also shown by the different behavior of the acoustic and elastic PP reflection coefficients versus angle beyond the critical incidence (Figure 6.3) (Červený *et al.*, 1977).

The radiation patterns of the V_S parameter for the last three diffraction modes (P-SV, SV-P and SV-SV) are shown in Figure 6.2(b-d). The union of these three radiation patterns covers the full range of aperture angles, with a higher sensitivity of the SV-SV mode to the V_S parameter. Therefore, a high resolution reconstruction of the V_S parameter is also expected if the four scattering modes can be exploited during FWI. The trade-off between the V_P and V_S parameters is expected to be limited because the radiation patterns of the two classes of parameters overlap only for the P-P mode at intermediate scattering angles. We shall assume that this parametrization analysis, which has been applied in the elastic isotropic approximation, will apply to VTI media. A detailed parametrization analysis of FWI for VTI media is left to future work.

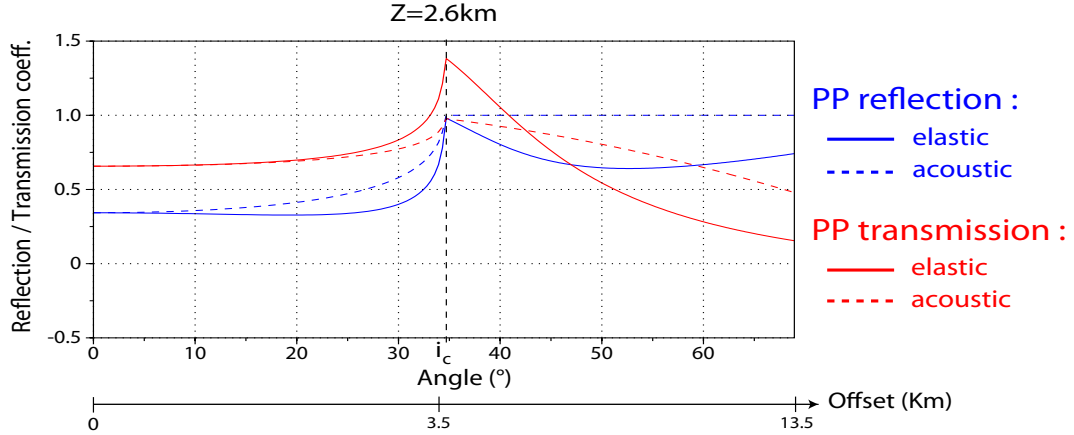


FIGURE 6.3 – P-P reflection (blue) and transmission (red) coefficients with respect to angle (and offset) at a liquid-liquid (dash curves) and solid-solid (solid curve) interface (Červený *et al.*, 1977). The interface is located at 2.6 km in depth. The P-wave velocities in the upper and lower medium are 1700 and 3000 m/s, the S-wave velocities are 650 and 1350 m/s and the densities are 2020 and 2340 kg/m^3 , respectively. These values are representative of the Valhall field. Note the differences between the acoustic and elastic AVO curves, beyond the critical distance of 3.5 km.

6.5.4 Starting models and inversion strategies for this study

The last inversion steps, presented in this study, consists of the joint update of the V_{P0} and V_S wavespeeds (Table 6.1). The initial V_{P0} model is the final V_{P0} model of the multiparameter visco-acoustic FWI presented in the Paper I (Figure 6.4(a,e)) and the initial V_S model was built by anisotropic reflection traveltime tomography and was provided by the BP company (Figure 6.4(b,f)). The background models of the density, attenuation and Thomsen's parameters are kept fixed to their initial values and are described in the Paper I. The update of the V_{P0} parameter during the first two steps provides a more accurate starting model of V_{P0} to update the V_S parameter from either the P-P or the P-S reflections during the late inversion steps. Of note, the possibility to perform a hierarchical update of the V_{P0} and V_S parameters is a clear advantage of marine data relatively to onshore data, where the strong sensitivity of the surface waves to the two wavespeeds may require to update simultaneously these two classes of parameter (Brossier *et al.*, 2009b).

In this study, we shall test two main data-driven strategies for the joint update of V_{P0} and V_S (Table 6.1) : in the first strategy, the hydrophone data are first inverted for updating the long wavelengths of the shear wavespeed using the amplitude-versus-offset variations of the P-P reflections (Sears *et al.*, 2008, 2010) (referred to as Test H in the following). In a second step, the geophone data are inverted to update the short-wavelengths of both the vertical compressional wavespeed and the shear wavespeeds (referred to as Test HG in the following). In the second strategy, only the geophone data are inverted to jointly update V_{P0} and V_S (referred to as Test G in the following).

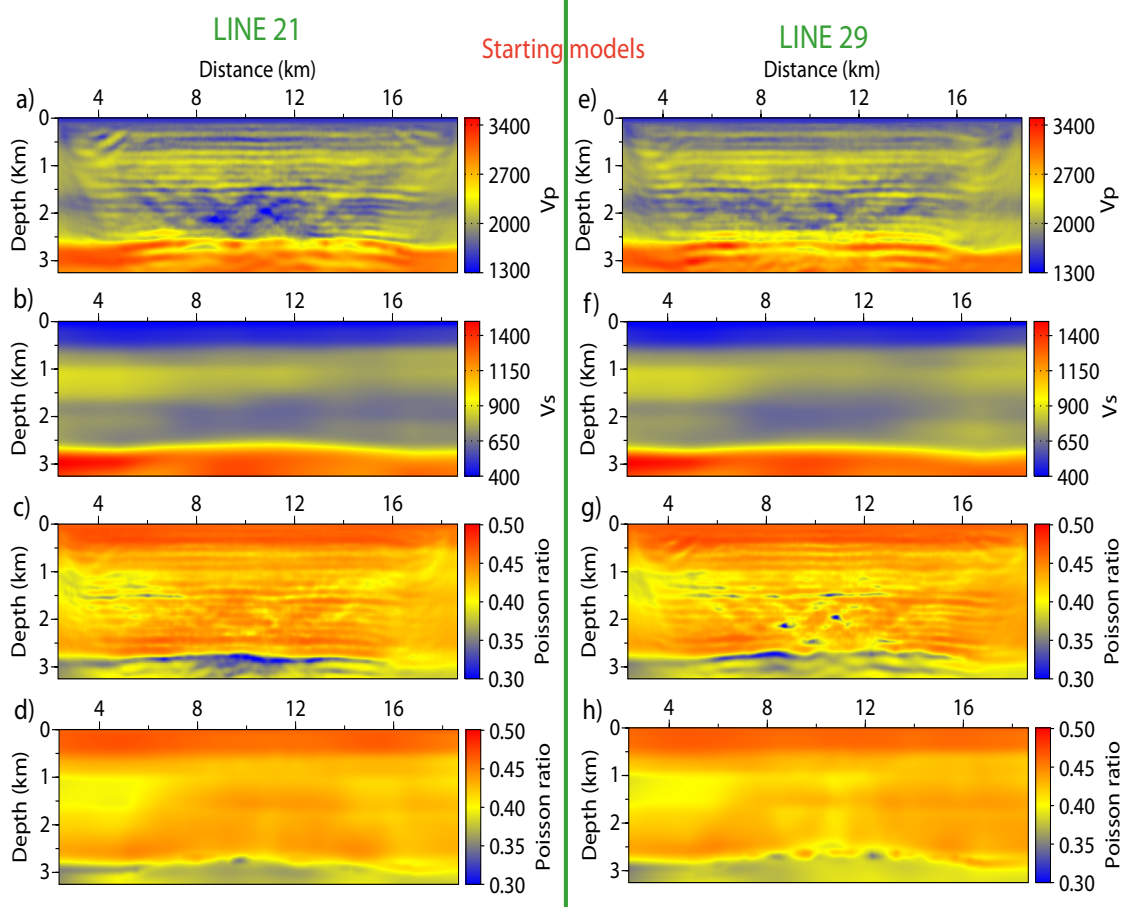


FIGURE 6.4 – Starting models of elastic FWI of line 21(a-c) and line 29 (e-g). (a, e) The P-wave velocity model is inferred from acoustic FWI (a,e). The S-wave velocity model was built by anisotropic reflection traveltime tomography. (c, g) Poisson's ratio models inferred from the P-wave and S-wave velocity models shown in (a-b) and (e-f), respectively. (d-h) Same as (c,g) except that the P-wave velocity model used to compute the Poisson's ration was built by reflection traveltime tomography instead of acoustic FWI.

6.6 Data and FWI setting

The dataset consists of 220 multi-component receivers on the sea bottom and 320 explosive shots, located 6 m below the sea surface. Two cables from the 3D OBC survey (21 and 29) are processed in this study (Figure 14 of Paper I). The cable 21 is located near the end of the gas cloud imaged in Sirgue *et al.* (2010), while the cable 29 is located off the gas cloud. The maximum source-receiver offset is 13 km. Following the same strategy than for the visco-acoustic FWI presented in the Paper I, we exploit the reciprocity holding between a source and a receiver to save computational cost during visco-elastic inversion of the hydrophone component. In contrast, the reciprocity is not used during the inversion of the geophone components because it would require to perform two simulations per vertical and horizontal geophones (treated as vertical and horizontal forces), leading to a total number of 440 simulations to compute the

incident wavefields instead of 320 when the reciprocity is not used. The numerical cost of the elastic modeling is about one order of magnitude higher than in the acoustic case because of very low shear-wave velocities at the sea bottom, which are four times smaller than the smallest V_{P0} value. This leads to a number of cells in the elastic mesh, which is about sixteen times higher than in the acoustic case. The maximum number of cells reaches a value of $4.7 \cdot 10^6$ at 6.7 Hz, that prompts us to adapt the mesh size to the frequency.

During FWI, we consider the five following groups of 3 overlapping frequencies when we invert the hydrophone data : $f_1 = [3.5, 3.78, 4]$, $f_2 = [4, 4.3, 4.64]$, $f_3 = [4.64, 5, 5.25]$, $f_4 = [5.25, 5.6, 6]$, $f_5 = [6, 6.35, 6.7]Hz$, while we consider only the f_2 , f_3 , f_4 and f_5 frequency groups for the geophone-data inversion (renamed f_1 , f_2 , f_3 and f_4 in Figures 6.5(a-d,i-l) and 6.5(a-b,e-f)), as the signal to noise ratio of the geophone data is considered too poor at 3.5 Hz. The recorded data are transformed in the frequency domain considering 4 s and 8 s of signal for the hydrophone and the geophone components respectively.

The FWI geophone data pre-processing consists in a minimum-phase whitening, followed by Butterworth filtering with cut-off frequencies of 3 and 20 Hz, a mute before the first arrival, and a 3D to 2D amplitude correction implemented by multiplying the data by \sqrt{t} . No time damping is applied to the hydrophone and geophone data during the inversion and, hence we seek to use the full information content of the data components in one go. Moreover, no gain with offset is applied to the data during inversion ($\mathbf{W}_d = \mathbf{I}$ in equation 6.1). The L-BFGS optimization performed less than 12 iterations per frequency group, although the maximum number of iterations was set to 25. A source wavelet per shot gather is estimated at each iteration of the hydrophone and geophone data by solving a linear inverse problem (Pratt, 1999, his equation 17).

The standard errors σ_i , which are used in the roughness matrices \mathbf{W}_{m_i} , equation 6.2, are set to 10 % of the mean value of the parameter class i . The damping factors λ_i are defined according to the sensitivity of the data to the parameter classes, which can be assessed from the diagonal terms of the approximate Hessian, equation 6.5, namely, the auto-correlation of the partial derivative wavefields at the receiver position. A suitable tuning of the damping terms λ_i must be chosen such that the scaling of the gradients of the misfit function performed by the Hessian remains effective, while the conditioning of the Hessian matrix is sufficiently-high to guarantee a stable reconstruction of all of the parameters. Of note, if the sensitivity of the data to a parameter class is too weak and the signal-to-noise ratio in the data is too poor, it is unlikely that the inversion will succeed in imaging the model parameter, whatever the value of λ_i . As for the multiparameter acoustic inversions presented in the Paper I, our choice of the damping parameter leads to a data gradient 10 times higher than the model gradient. The maximum value of the data Hessian for V_{P0} is in average four times higher than the data Hessian for V_S for all of the tests, and therefore, we use a damping value 4 times smaller for V_S than the one used for V_{P0} . As we shall see, we have verified that the updated V_{P0} and V_S models show a reasonably-good agreement with sonic logs for this choice of dampings. The standard error and the damping values used for the V_{P0} and V_S parameter classes are outlined in Table 6.2.

Parameter	σ	λ
V_{P0}	200 (m/s)	4.10^{-8}
ρ	200 (kg/m ³)	4.10^{-8}
Q_P	15	4.10^{-10}
V_S	100 (m/s)	1.10^{-8}

TABLE 6.2 – Table synthesizing the values taken by the uncertainties σ and the inversion dampings λ for each inverted class of parameter.

6.7 FWI results

The elastic FWI models obtained with the Tests G, H and HG are shown for lines 29 and 21 in Figures 6.5 and 6.6, respectively. The starting and the final FWI V_{P0} models (Figures 6.4(a,e), 6.5(d,h,l) and 6.6(a,c,e)) do not show significant differences, these differences reaching a maximum value of ± 250 m/s. Limited update of the V_{P0} parameter was expected as the former inversion of the hydrophone data in the visco-acoustic approximation has yet provided a reliable V_{P0} model. The progressive increase of resolution of the V_S FWI models as the inversion progresses towards higher frequencies is highlighted in Figure 6.5. The V_S model inferred from the hydrophone component during the first step of the strategy 1 (Test H) is smoother than the one inferred from the geophone components during the second step (Test HG), that is consistent with the radiation pattern analysis described in the previous section. The final FWI V_S model inferred from the hierarchical inversion of the hydrophone and geophone data (Test HG) shows thin continuous reflectors at 0.6 km, 1.25, 1.5 km and 2.4 km in depth. These reflectors are better imaged when the hierarchical inversion of the hydrophone and geophone data (Test HG) is performed relatively to the inversion of the geophone components only (Test G). This highlights the benefit of the data-driven hierarchical approach over data component, that is used in this study.

6.8 Model appraisal

6.8.1 Comparison with sonic logs

We first assess the FWI V_S models against sonic logs (Figure 6.7). As already noticed, the amplitudes of the V_S contrasts are weaker when only the geophone data are inverted (Test G) relatively to the hierarchical inversion of the hydrophone (Test H) and geophone data (Test HG). The velocity contrasts are also weaker in the V_S model of line 21 relatively to the one of line 29. This may be related to the platform noise, which affects more the cable 21, the platform being closer to this cable. The starting V_S model shows overestimated velocities in the gas relatively to the sonic log. The full waveform inversion fails to update the long wavelengths of the V_S model to match the trend of the sonic log in the gas : mainly short-wavelength perturbations were added by FWI to match the small-scale structures of the sonic logs. This suggests that there is no information in the data to update the long to intermediate wavelengths of the V_S model. This lack of constraints on the long wavelengths of the S-wave velocity structure can be explained by the shorter propagated S wavelengths relatively to the P wavelengths and by the

LINE 29

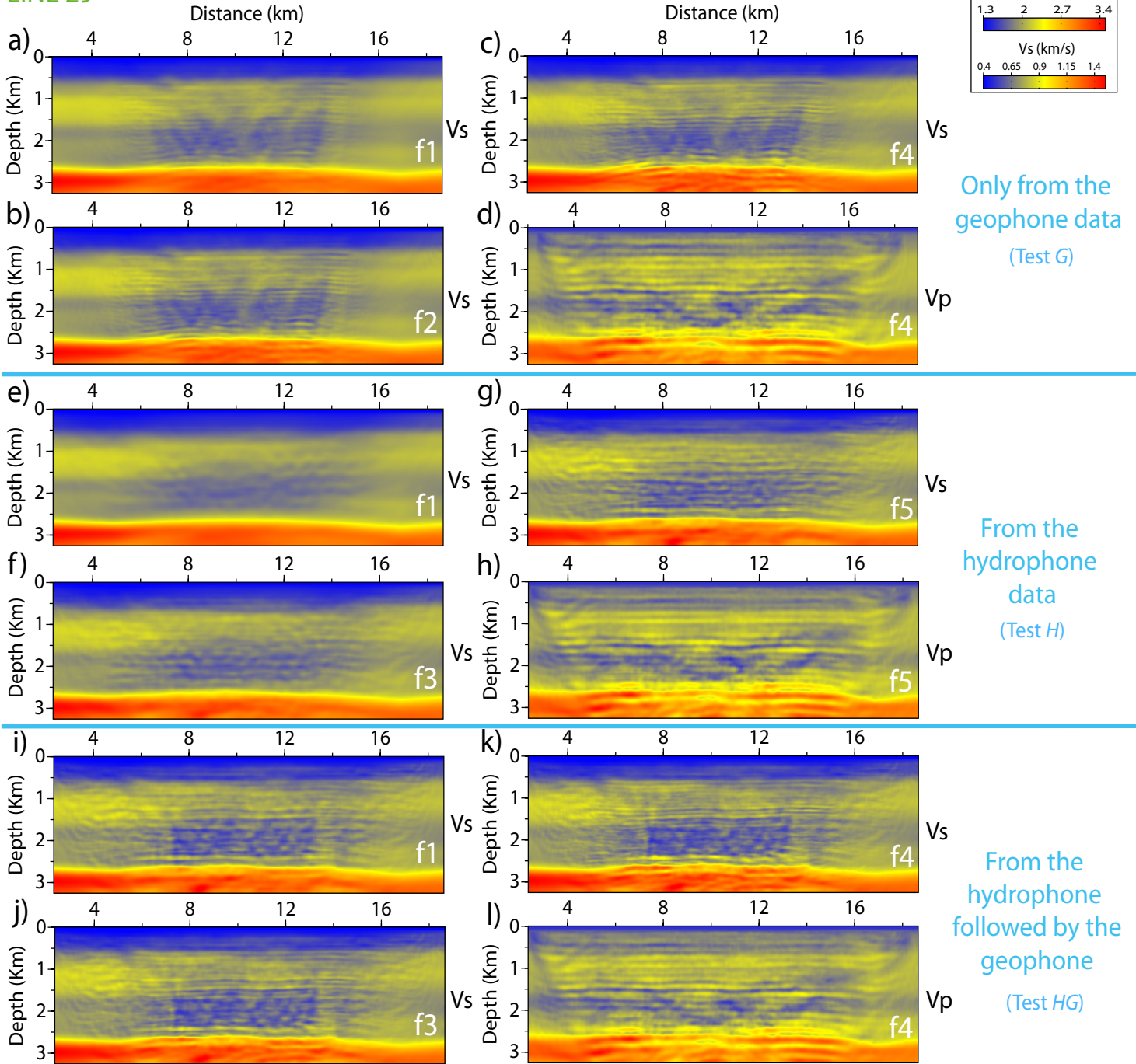


FIGURE 6.5 – Intermediate and final P-wave and S-wave velocity FWI models of Tests *G* (a-d), *H* (e-h) and *HG* (i-l) for line 29. Index of the different group of frequencies are denoted by f_i , the corresponding frequencies being given in the text.

narrow range of aperture angles over which P-S converted waves are recorded. The short-scale features, which show a good agreement between the FWI profile and the sonic log (Figure 6.7,

LINE 21

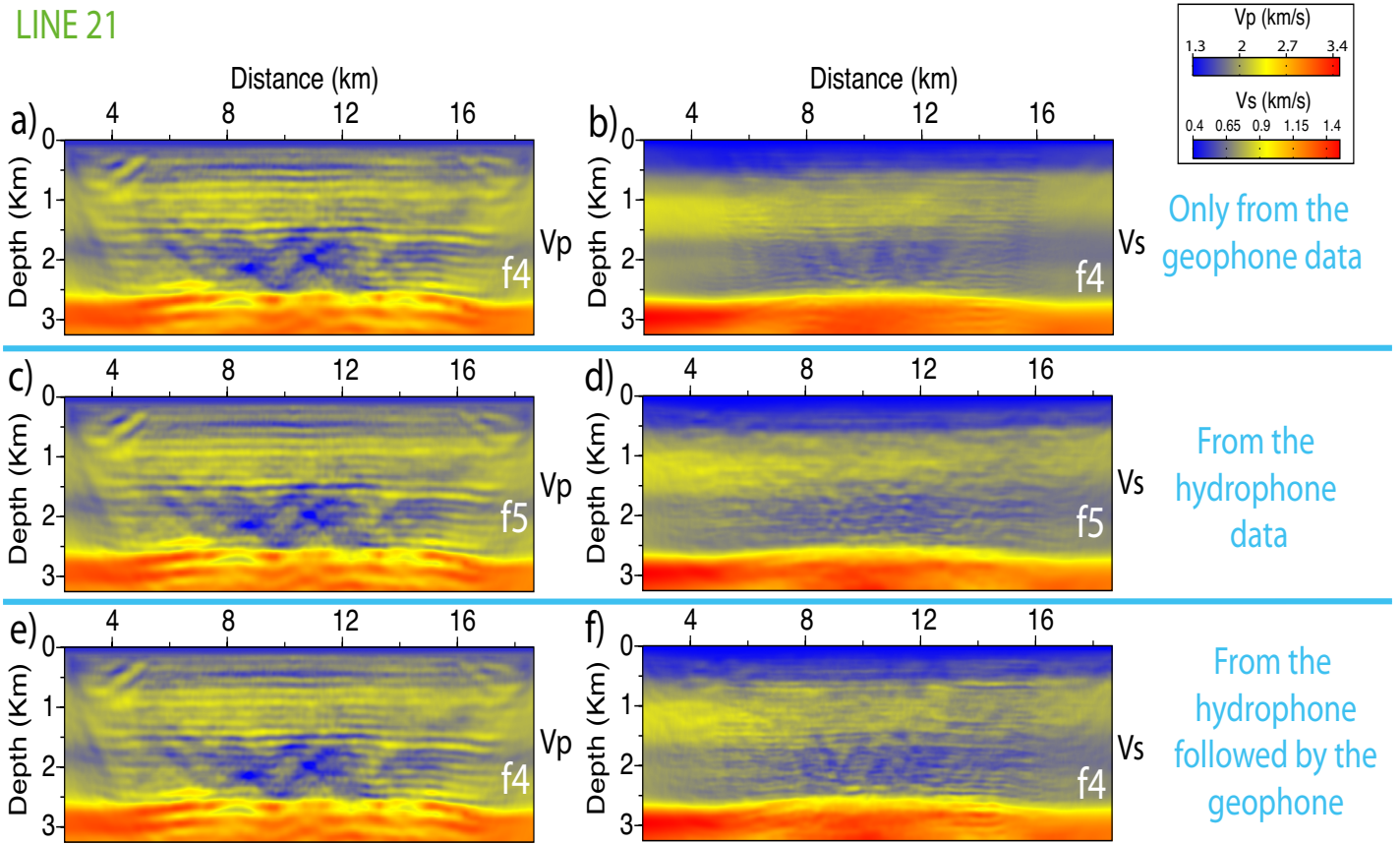


FIGURE 6.6 – Final P-wave velocity (left panel) and S-wave velocity (right panel) FWI models of Tests *H* (a-b), *G* (c-d) and *HG* (e-f) for line 21. Results appear noisier than for line 29 (Figure 6.5).

black dashed circles), mostly correspond to the continuous interfaces at 0.6 km, 1.5 km and 2.4 km in depth shown in Figure 6.5.

6.8.2 Data match

6.8.2.1 Anatomy of the data

An important assessment of the FWI results relies on the comparison between the recorded seismograms and the seismograms computed in the FWI models. Before showing this comparison, it is worth analyzing the imprint of the elastic properties of the subsurface in the data. The recorded horizontal and vertical geophone components of line 29 are shown in Figures 6.8a and 6.9a, on which are superimposed traveltime curves of the different type of converted waves computed by ray tracing in the FWI models of line 29 from Test *HG*. As the ray tracing that we used is isotropic, we stretch vertically the FWI model to roughly account for anisotropic effects at short offsets. Arrivals labeled in the seismograms of Figures 6.8a can be interpreted by examining few snapshots of the wavefronts of the horizontal particle velocities, computed in the final FWI model of Test *HG* with a VTI elastic finite difference time domain engine (Figure

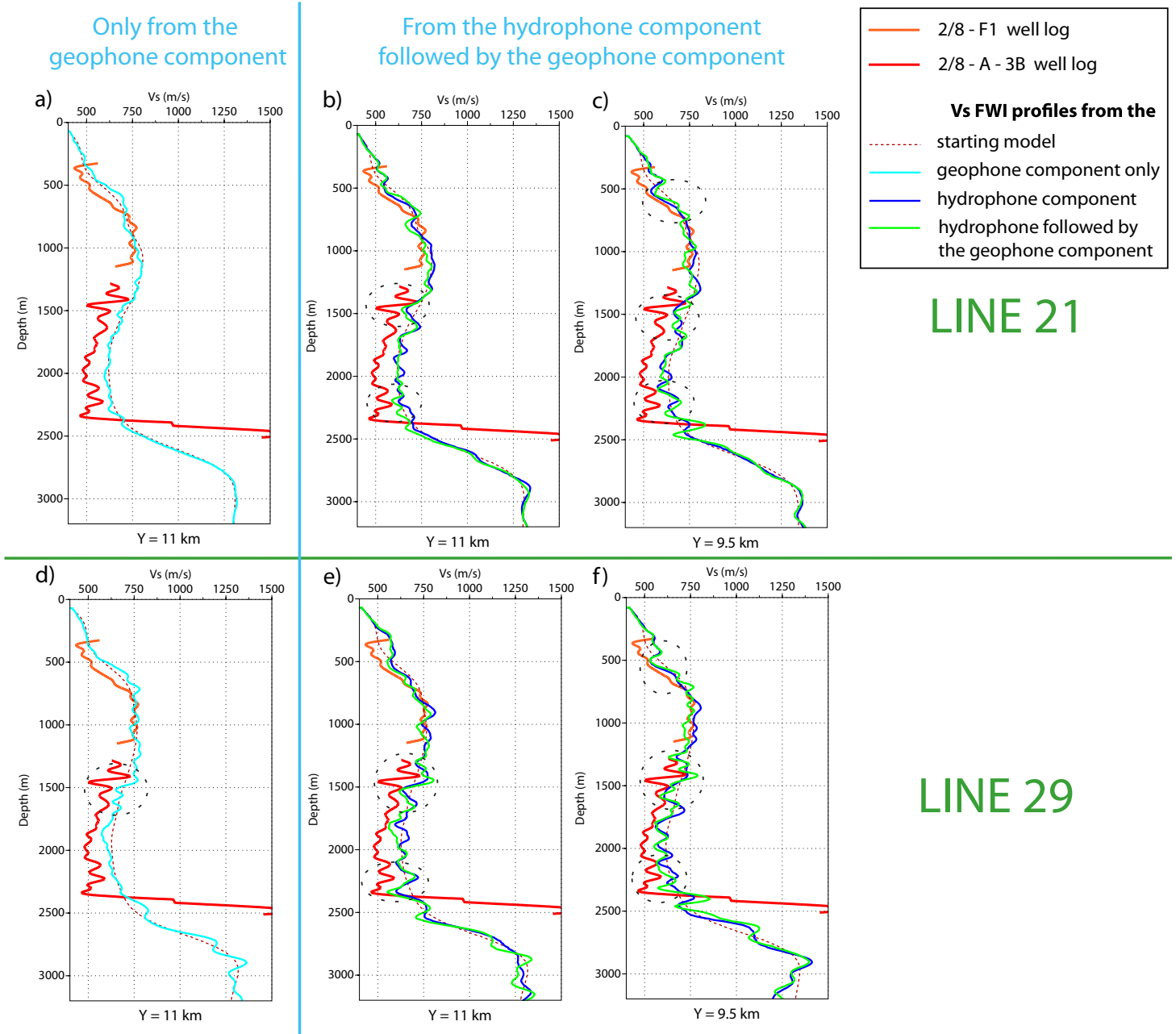


FIGURE 6.7 – Comparison between two sonic logs (red and orange curves) and the profiles of the final S-wave velocity FWI model of line 21 (a-c) and line 29 (d-f). The FWI profile is extracted at the well log position ($Y = 11$ km), and also at $Y = 9.5$ km for the hierarchical strategy. The results obtained from Test G (a,d), H and HG (b-c,e-f) are denoted by the cyan, the blue and the green curves, respectively. Location of the well logs is shown in Figure 14 of the companion paper.

6.10). The PP, PS and SS converted waves reflecting from the four main interfaces at 0.6 km, 1.25 km, 1.5 km and 2.6 km in depth can be clearly tracked in these snapshots. The traveltimes

curves of the P and S diving waves present a slope break at a cross-over distance of 3 km and 2 km respectively, corresponding to the distance beyond which the diving waves propagating between 0.6 and 1.5 km in depth are recorded as a first arrival. We do not see any evidence of S-S waves in the seismograms recorded by the horizontal geophone, although S-S reflections from reflectors at 1.25 km and 1.5 km in depth can be interpreted in the snapshots at times 3.6 s and 5.2 s with a non negligible amplitude, and are visible in the seismograms computed in final FWI model of Test *HG* (Figure 6.8c).

Amplitudes of P-S converted phases generally increase with depth as the shear velocity becomes closer to the compressional velocity (White & Stephen, 1980), that results from the rapid increase of V_S with depth (Hoven *et al.*, 1991). The V_P over V_S ratio in the final FWI model of Test *HG* is about 4 in the first 220 m in depth, 3 down to 2.5 km in depth, and 2 below 2.5 km in depth. This explains the weak PS conversion at the sea bottom, and why the strongest PS conversion occurs at the reservoir level, a ratio of 3 in the Tertiary sediments being rather high (Mueller *et al.*, 1997).

6.8.2.2 Seismic modeling

To compare recorded and modeled seismograms, we compute impulsional seismograms using a Dirac source wavelet within the [1-10] Hz frequency range and we convolve them with the source wavelet estimated from the corresponding recorded data component following the expression of Pratt (1999). We first compare horizontal-particle velocity seismograms computed in the FWI starting models of line 29 with the acoustic and the elastic approximations, respectively, to highlight the imprint of the shear wave velocity in the elastic wavefields (Figure 6.8(d-e)). The main differences between the two sets of seismograms represent the S diving waves, and the PS waves with a conversion at 2.6 km in depth. The main arrivals shown in both gathers are the PP arrivals reflecting from the top and the bottom of the gas layers. However, the $PS_{0.6}$, $PS_{1.25}$ and $PS_{1.5}$ converted waves are not visible. The elastic seismograms computed in the final FWI models of Test *H* show an enhanced $PS_{2.6}$ arrival and a weak $PS_{1.5}$ wave (Figure 6.8b). All the PS waves shown in the recorded gather can be observed in the seismograms computed in the final FWI model of Test *HG*, hence supporting that FWI successfully extracted this information (Figure 6.8c). Converted P-S arrivals in the seismograms computed in the final FWI model of Test *G* (Figure 6.8f) have smaller amplitudes than those computed in the final FWI model of Test *HG* (Figure 6.8c), that supports the superiority of the hierarchical approach by successive inversions of the hydrophone and of the geophone components. A more detailed view of the waveform agreement between recorded and modeled data is provided by the direct comparison of the two sets of seismograms plotted with their true amplitudes in Figure 6.11.

The seismograms recorded by the vertical geophones of line 29 (Figure 6.9a) do not show any evidence of PS and SS converted waves. The V_Z seismograms computed in the starting and in the final FWI models of Test *HG* (Figure 6.9(b-c)) show the PP reflections from the top and the bottom of the gas. The amplitudes of these reflections have been however enhanced after FWI, in particular at short offsets (Figure 6.9c). It is worth noting that the short-offset reflections in the pressure seismograms computed in the final FWI model of Test *HG* (Figure 6.9f) have weaker amplitudes than those computed in the V_{P0} starting model with the acoustic approximation (Figure 6.9e). This shows the non-negligible effect of the elastic subsurface

properties on the AVO of the PP reflections. The reflection arriving at a two-way travelttime of 1 s at zero-offset, which corresponds to the reflection from the interface at 0.6 km in depth, has also been improved in the seismograms computed in the final FWI model (Figure 6.9f, Test HG). The direct comparison between the recorded and modeled seismograms for the pressure and V_Z components is shown in Figure 6.12.

The seismograms recorded by the horizontal geophones of line 21 clearly show the platform noise of hyperbolic shape (Figure 6.13a). The PS phases have stronger amplitudes in the shot gather computed in the FWI model of Test *HG* (Figure 6.13e and 6.11e) than in the one computed in the FWI model of Test *G* (Figure 6.13c), although these amplitudes remain weaker relatively to those shown for line 29 (Figure 6.8c). It is interesting to note that the $PS_{2,6}$ arrival is not visible in the seismograms computed in all of the FWI models of line 21 between -4 km and -5 km in offset. This range of offset corresponds to the location of the platform noise in the recorded data, hence suggesting that the inversion was significantly hampered by this coherent noise.

6.8.3 Source wavelet estimation

Another way to appraise the models is to estimate a source wavelet per shot gather by solving a linear inverse problem following the approach of Pratt (1999) and to test its repeatability from one shot to the next (Jaiswal *et al.*, 2009; Malinowski *et al.*, 2011; Prioux *et al.*, 2011). The peak-to-peak amplitude of the mean wavelet stacked over all the shots is representative of the repeatability of the estimated wavelets. The mean wavelet estimated from each data component and from the initial and final FWI models are shown in Figure 6.14. For both lines 21 and 29 and regardless the component, we observe the progressive improvement of the peak-to-peak mean wavelet amplitude from the starting model to the final FWI models of Test *H* and Test *HG*. The mean wavelet amplitude inferred from the FWI model of Test *HG* and from the hydrophone component is exactly the same than the one inferred from the FWI model of Test *H* because the pressure is not inverted during Test *HG*. We also notice that the mean wavelets inferred from the FWI model of Test *G* have always a smaller amplitude than the ones inferred from the FWI models of Test *HG*. One can note that the mean wavelets inferred from the two geophone components are very similar and differ from the wavelet inferred from the hydrophone. We can remind that, during the elastic inversion of the geophone data, the source is estimated from the two particle-velocity components. We show in Figure 6.15 all of the estimated wavelets per shot, that are computed in the starting and final FWI models of Test *HG* for all of the data components. We show rather repeatable wavelets, except near the ends of the acquisition where short offsets are lacking. The increase of amplitude in the final model is mainly seen for the shots located near the ends of the acquisition. The improvement of the wavelet focusing achieved in this study after elastic FWI is less significant than in previous acoustic FWI applications (Malinowski *et al.*, 2011; Prioux *et al.*, 2011). This results because the starting V_{P0} model used in this study for elastic FWI has been built by visco-acoustic FWI during an earlier stage and accounts for a large part of the elastic wavefield. Therefore, the improvement of the wavelet focusing shown in this study mainly highlights the secondary contribution of the S-wave velocity structure to match the elastic wavefield in most marine environments. The footprint of the platform noise is also clearly visible in the plot of the wavelets of line 21 estimated from the V_X component. Some noise is shown at the wavelet index 175, which corresponds to a in-line position of 11.25 km, that is the platform location.

LINE 29 - Vx component - $X_S = 12.5$ km

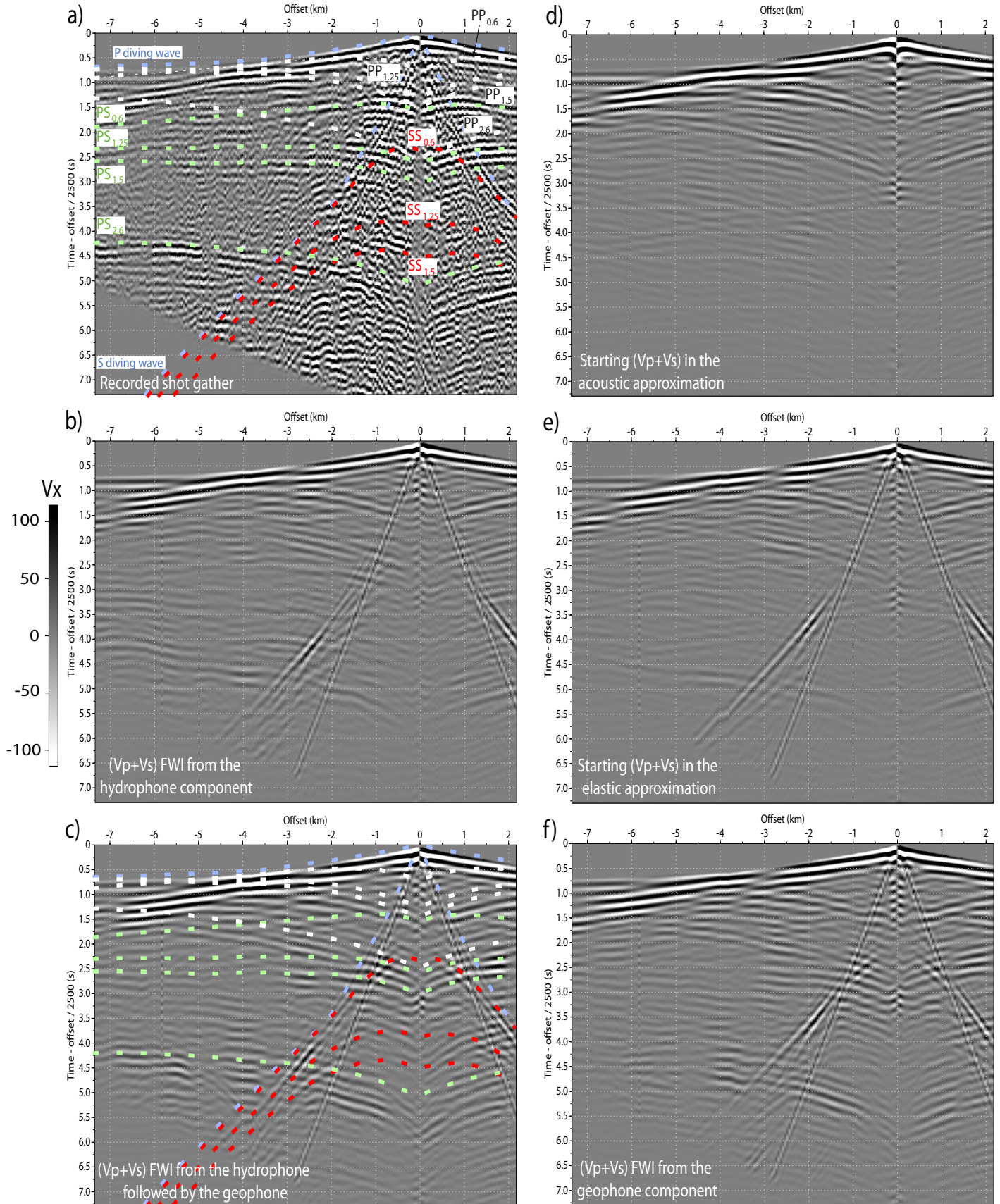


FIGURE 6.8 – Shot gather recorded by the horizontal geophones of cable 29. The shot is located at $X_S = 12.5$ km. (a) Recorded data, on which traveltimes of the various converted waves computed by ray tracing in the final FWI model of Test HG are superimposed. (b, c, f) Computed data in the final models of Test *H* (b), *HG* (c) and *G* (f). (d-e) Seismograms computed in the initial models with the acoustic (d) and elastic (e) approximation can be compared to assess the imprint of the S-wave velocity structure. PP, PS and SS denote a P-P, a P-S and a S-S reflection, respectively (the propagation in the water layer is not taken into account in the nomenclature). The index denotes the depth at which the reflection takes place.

LINE 29 - $X_S = 12.5$ km

Vz component

P component

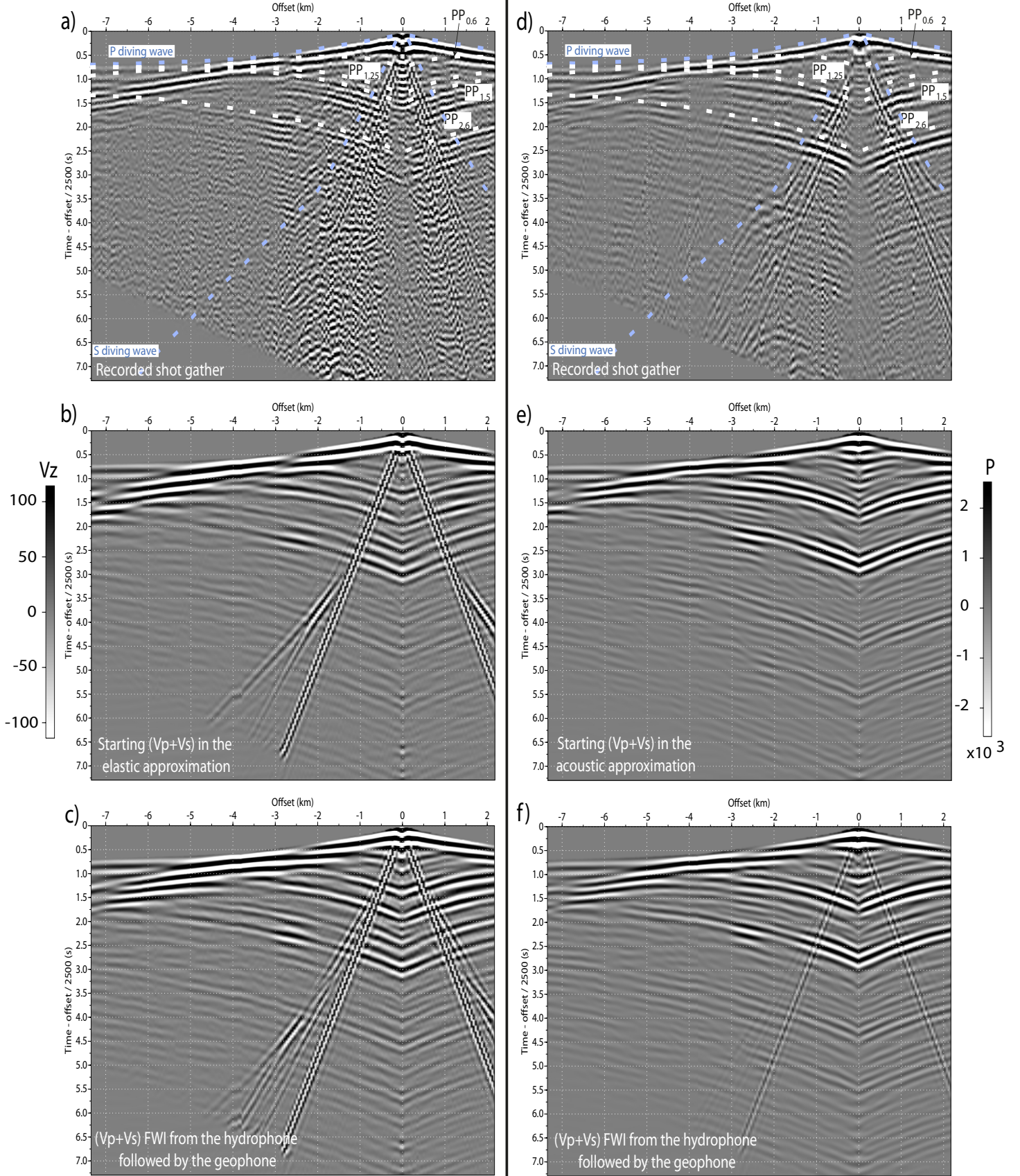


FIGURE 6.9 – Shot gather recorded by the vertical geophones (left panels) and the hydrophones (right panels) of cable 29. The shot is located at $X_S = 12.5$ km. (a,d) Recorded shot gather. (b,e) Synthetic seismograms computed in the initial models for the vertical geophone component in the elastic approximation (b) and for the hydrophone with the acoustic approximation (e). (c,f) Synthetic seismograms computed in the final models of Test *HG*. The same nomenclature than in Figure 7 is used.

LINE 29 - Vx component

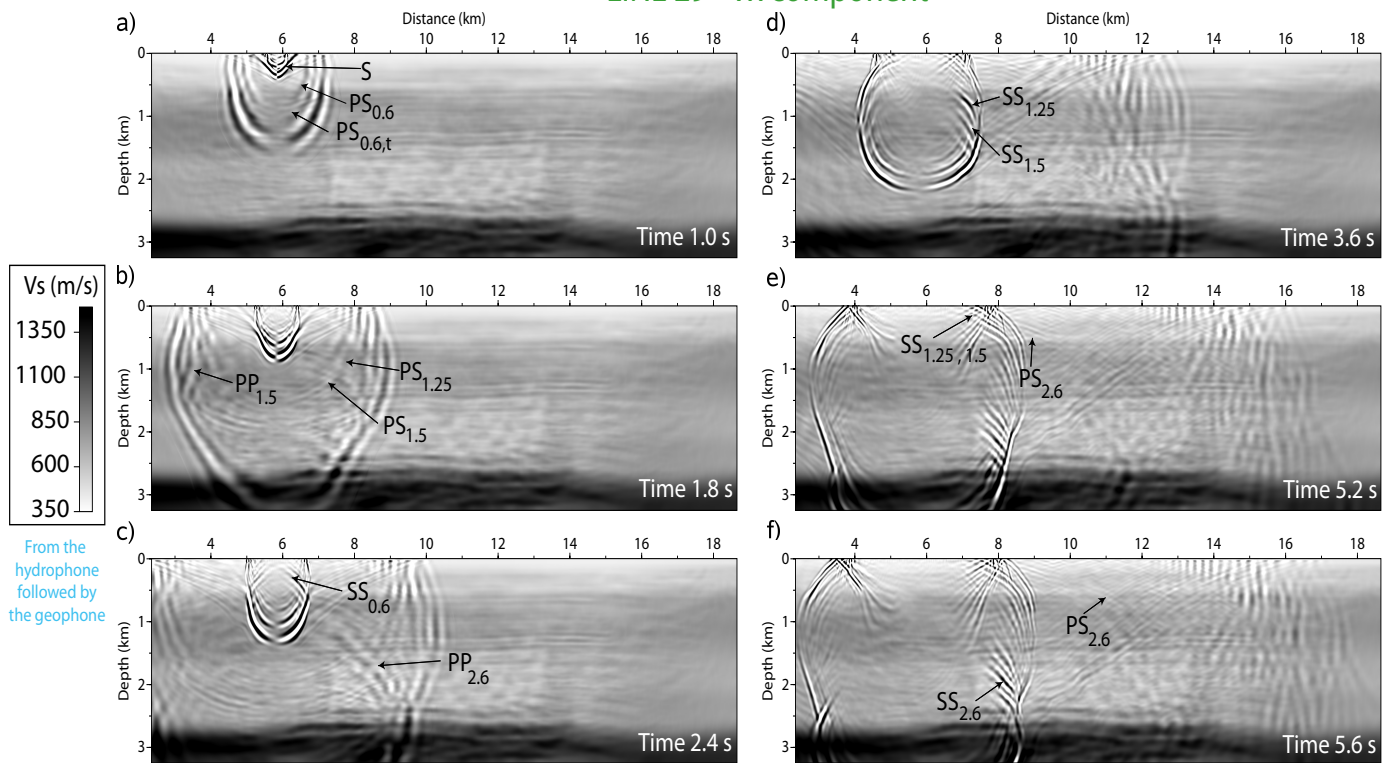


FIGURE 6.10 – Snapshots of elastic wavefields computed and superimposed on the final S-wave velocity FWI model of Test HG. The same nomenclature than in Figure 7 is used.

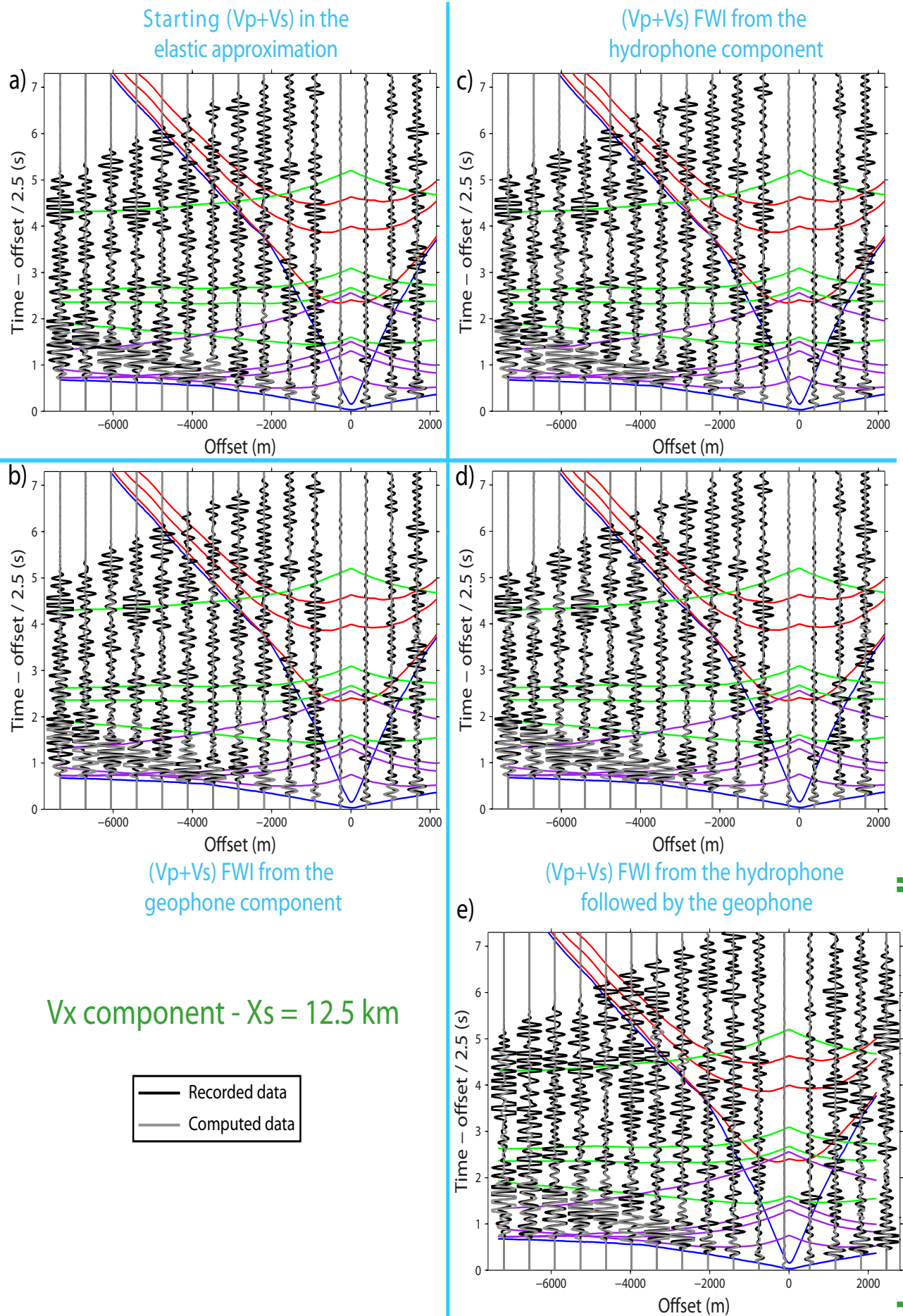


FIGURE 6.11 – Direct comparison between the observed (black line) and the computed seismograms (gray line) of line 29 (a-d) and of line 21 (e) for the horizontal-geophone component. The synthetic seismograms are computed in the starting models (a) and in the final FWI models of Test *G* (b), *H* (c) and *HG* (d-e).

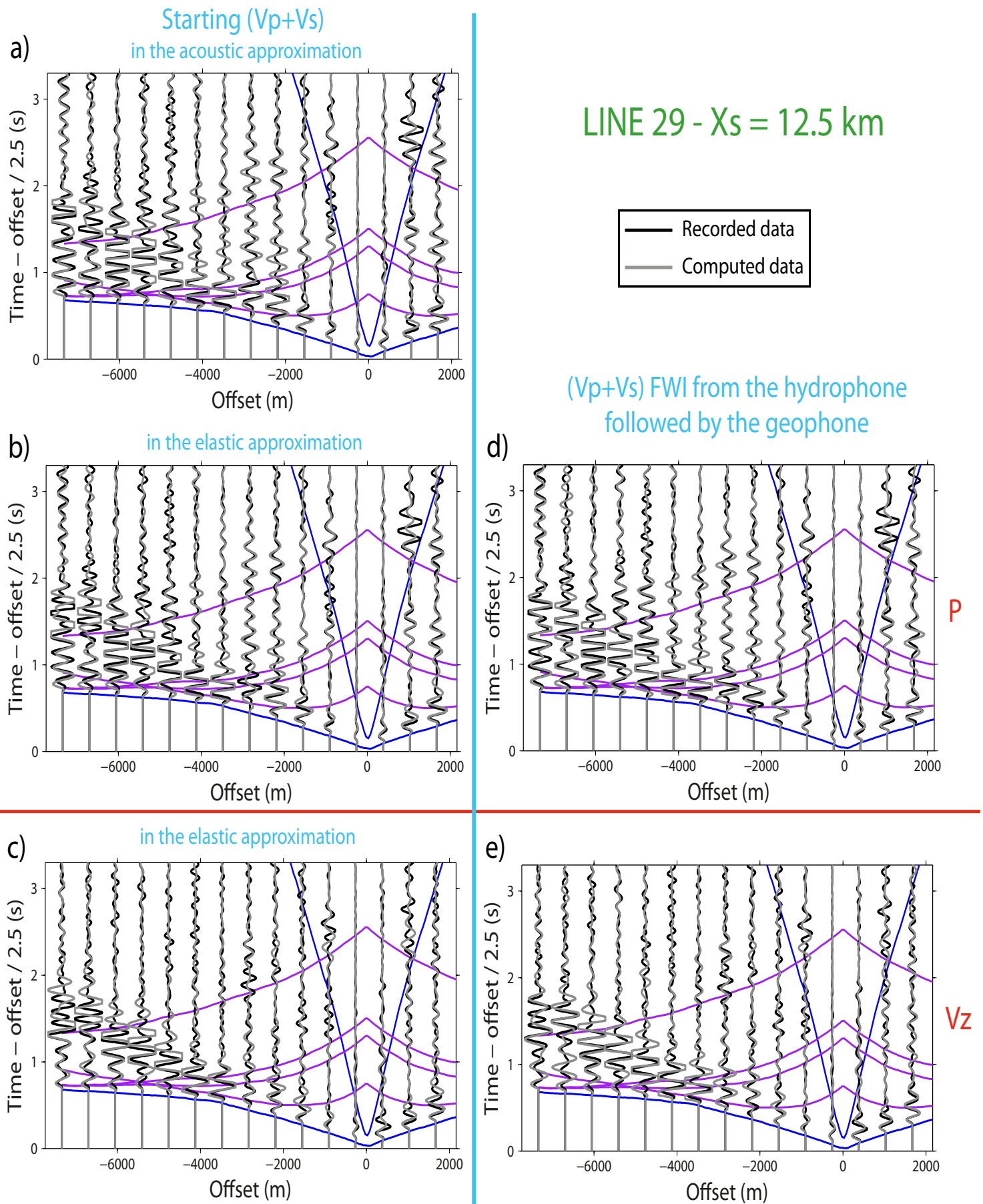


FIGURE 6.12 – Direct comparison between the observed (black line) and the computed seismograms (gray line) of line 29 for the hydrophone (a,b,d) and the vertical geophone (c,e). The pressure synthetic seismograms are computed in the starting models with the acoustic (a) and elastic (b) approximations and in the final model of Test HG with the elastic approximation (d). The particle-velocity seismograms are computed in the initial models (c) and in the final FWI models (e) of Test HG.

LINE 21 - Vx component - $X_s = 15.5$ km

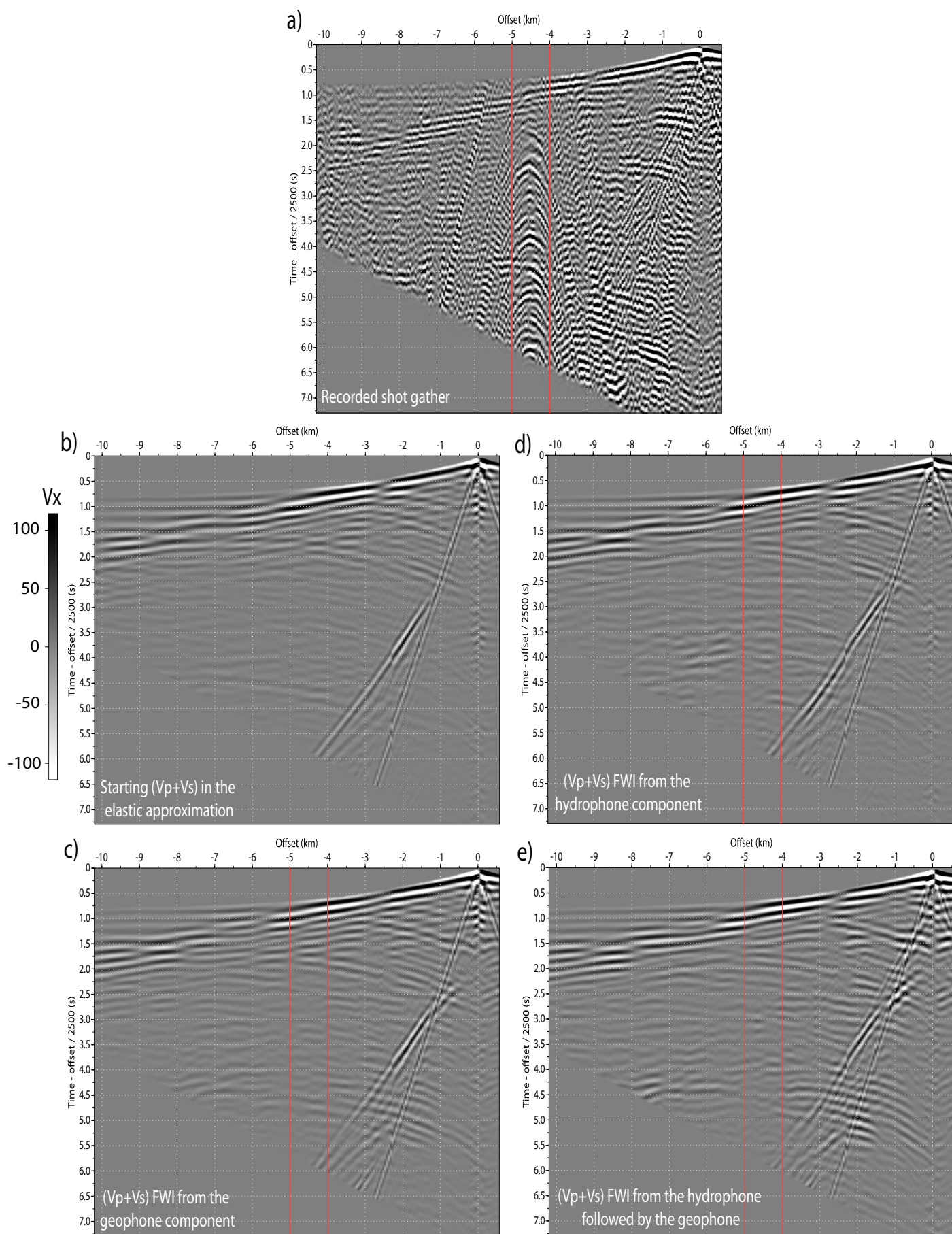


FIGURE 6.13 – Shot gather recorded by the horizontal geophones of line 21. The shot is located at $x = 15.5$ km of distance. (a) Recorded data. (b-e) Seismograms computed in the starting models (b), and in the final FWI models of Test *G* (c), *H* (d) and *HG* (e). The offset range, for which the platform noise pollute the data, are delineated by the red lines.

LINE 29

LINE 21

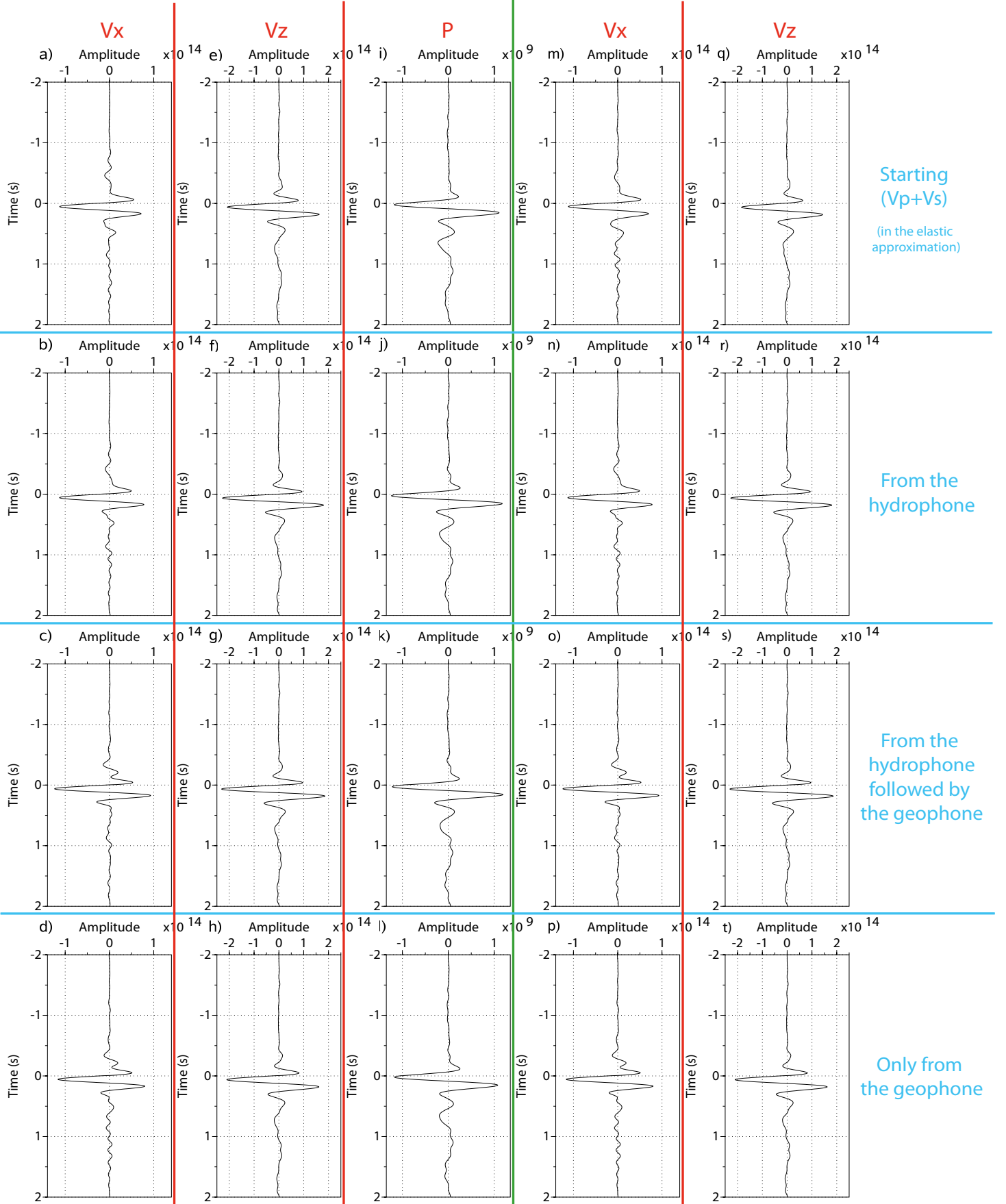


FIGURE 6.14 – Mean estimated source wavelets computed from the starting models (first row) and in the final models of Test *H* (second row), *HG* (third row) and *G* (fourth row). On each row, we show the wavelets estimated from the horizontal geophones, the vertical geophones and the hydrophones of line 29 (columns 1 to 3), and from the horizontal geophones and the vertical geophones of line 21 (columns 4 to 5).

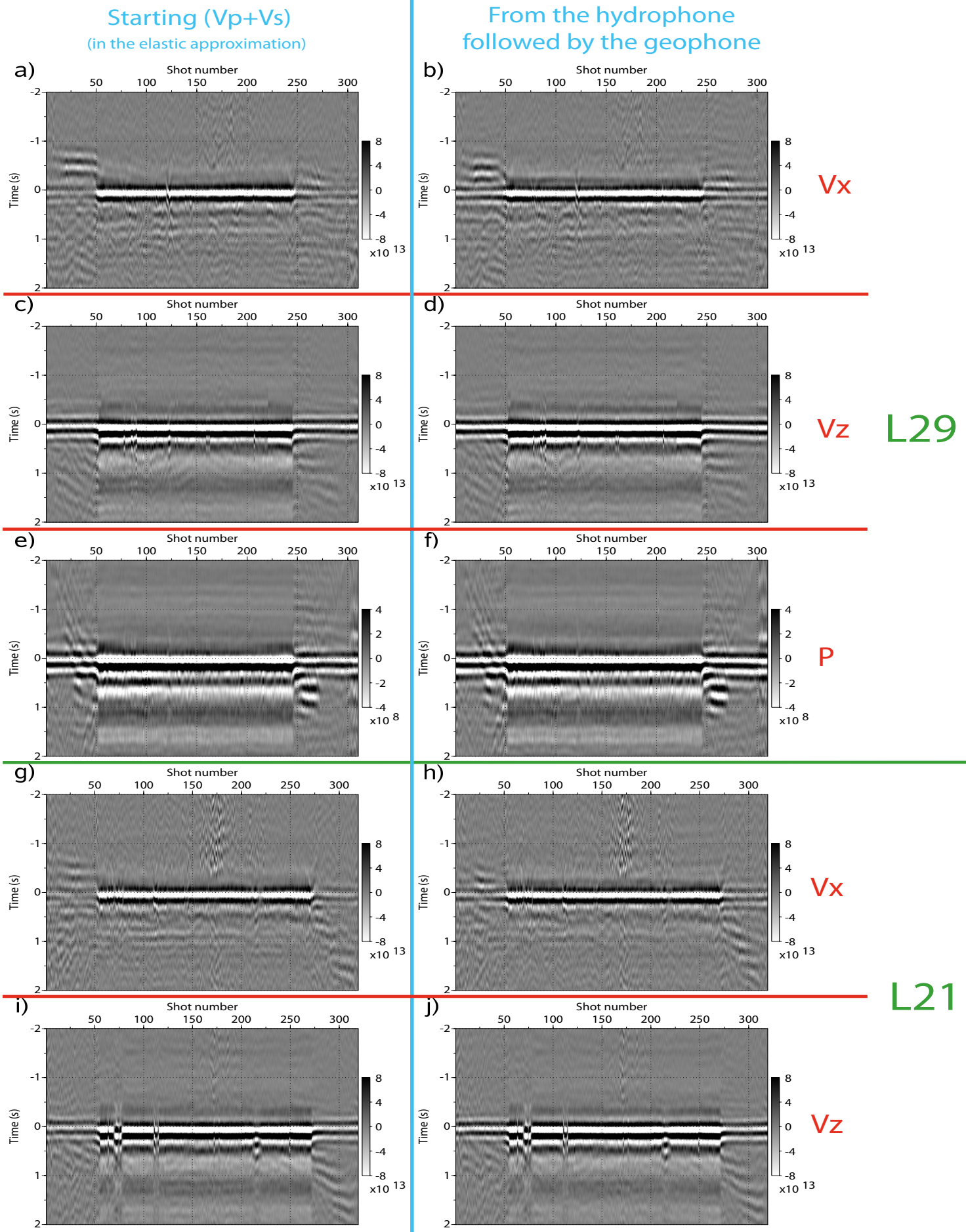


FIGURE 6.15 – Source wavelets estimated for each shot gather from the starting models (left panel) and the final FWI models of Test *HG* (right panel). We show the wavelets estimated for line 29 from the horizontal geophones (a-b), vertical geophones (c-d) and the hydrophones (e-f) of line 29, and from the horizontal geophones (g-h) and the vertical geophones (i-j) of line 21.

6.9 Derivation of related physical quantities and geological implications

6.9.1 Shallow shear-wave velocity structure

We found interesting to put in comparison the elastic starting models from lines 21 et 29 (Figure 6.16(a,c)) and the final V_S FWI models from Test *HG* (Figure 6.16(b,d)), with the Scholte wave velocity models inferred from passive seismic-noise tomography by Ridder & Dellinger (2011) and Mordret *et al.* (2011) (Figures 6.16(e-f)). The Scholte wave is a liquid solid interface wave whose velocity is less than V_S and is more strongly related to V_S than to V_P (Carcione & Helle, 2004). The geological features should thus affect in the same way the shear waves and the Scholte waves. The output of seismic-noise tomography is an horizontal cross-section, over the whole area covered by the receiver array, and represents the Scholte wave velocity within the first 200 m in depth. Although the Scholte wave velocity map from Ridder & Dellinger (2011) covers the whole receiver area and shows more details than the one from Mordret *et al.* (2011), both images are similar. In particular, the Scholte velocity is higher in the center of the field than at its periphery, that has been related to the subsidence of the tertiary sediments. This subsidence is caused by the compaction of the reservoir, and induce a contraction of the sediments at the center and an extension towards the edges of the field (Hatchell *et al.*, 2005). Furthermore, some thin high velocity anomalies in the velocity map of Ridder & Dellinger (2011) are interpreted as paleochannels (Figure 6.16e, red arrows), which can also be guessed in the top right of the map from Mordret *et al.* (2011). We find some striking similarities between our V_S FWI models and the Scholte wave velocities, although we acknowledge that the V_S starting model built by reflection traveltime tomography already describe the large spatial wavelengths of the velocity variations. The V_S FWI models present locally higher values around the positions $X = 10$ km and 13.5 km for line 21, and at the positions $X = 11.5$ km and more generally between 7.8 and 12.5 km for line 29. Subsidence effects are more visible for line 29, while line 21 seems more sensitive to the paleochannels, although the contrasts are not very sharp.

6.9.2 Poisson's ratio and other combinations of V_P and V_S

Interpretation of the vertical velocity, density and attenuation FWI models, developed in the Paper I, has suggested the presence of an accumulation of gas below barriers of claystones at 0.6 km and 1.5 km depth, and of two gas chimney for line 29. The elastic model developed in this study should help to confirm or negate our hypothesis as the joint knowledge of compressional and shear velocities is known to be of interest for fluid characterization. In rock physics, the Poisson's ratio is the ratio of strain normal to strain parallel to a uni-axial stress applied to a unit cube of rock. Limiting values are 0 (for no auxetic materials), when the parallel strain is solely accounted by a change in volume (no normal strain, such as a cork), and 0.5 when the normal strain entirely balances the parallel strain (no volumetric change, such as rubber). This quantity is thus related to compressibility, which increases as the Poisson ratio decreases. Presence of fluids in sediments tends to increase the compressibility, hence decreasing the

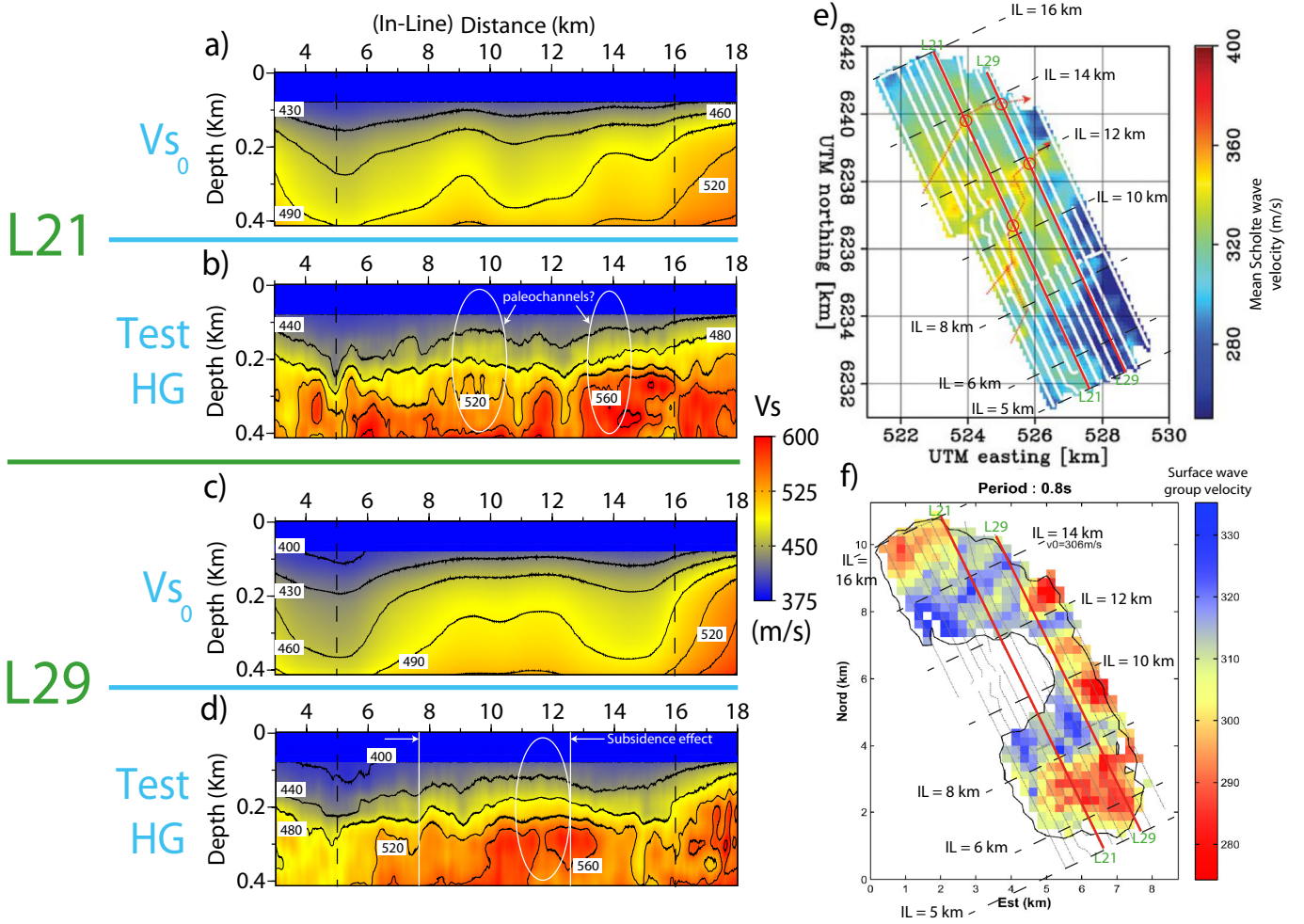


FIGURE 6.16 – Enlargements of the S-wave velocity starting models (a,c) and of the final FWI models of Test *HG* (b,d) for lines 21 (a-b) and 29 (c-d). (e) Mean Scholte wave velocity inferred by Ridder & Dellinger (2011) from ambient-seismic-noise eikonal tomography using 29 hours recording time in 2004, representative of the first 105 m below the sea floor. (f) Scholte wave group velocity at 0.8 s of period obtained by (Mordret *et al.*, 2011) from surface wave tomography, which is based on the inversion of dispersion curves constructed from the seismic noise correlation functions using 6 hours recording time.

Poisson ratio. In seismic, Poisson ratio (ν) is expressed as a function of V_P and V_S as

$$\nu = \frac{\frac{1}{2} \left(\frac{V_P}{V_S} \right)^2 - 1}{\left(\frac{V_P}{V_S} \right)^2 - 1}. \quad (6.7)$$

The shear-wave velocity is less affected than the compressional-wave velocity by the pore fluids as shear waves only propagate in the solid matrix. The compressional-wave velocity decreases in gas-charged sediments, while the shear-wave velocity does not change, hence de-

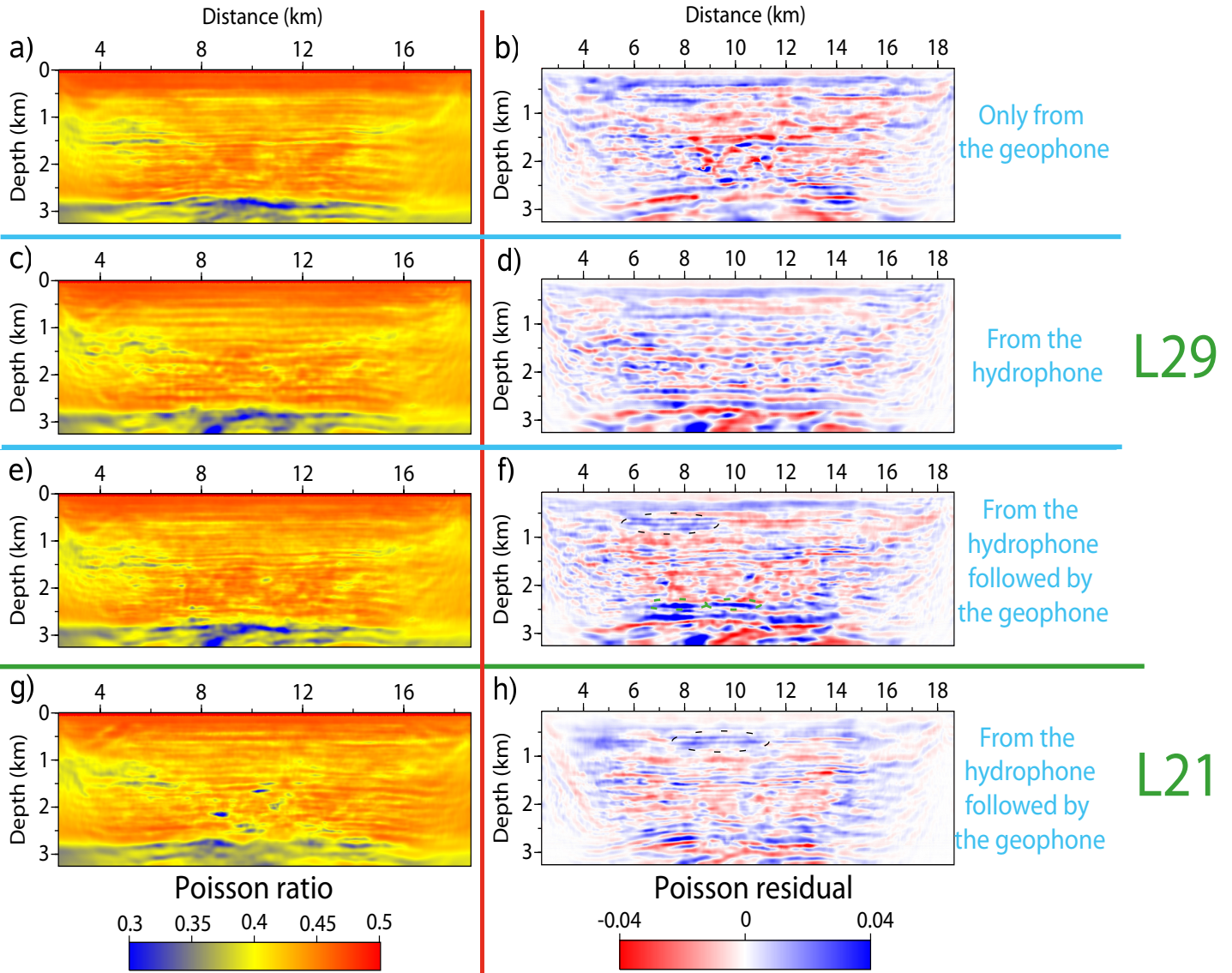


FIGURE 6.17 – Poisson's ratio derived using the final P-wave velocity and S-wave velocity models of line 29 inferred from Test *G* (a), *H* (c) and *HG* (e). (g) Same than (e) for line 21. (b,d,f,h) Residual Poisson's model formed by the difference between the starting models (Figure 6.4(c,g)) and the models shown in the left panel.

creasing the Poisson's ratio value. We show in Figure 6.17 the Poisson's ratio models inferred from the different inversion tests, and the corresponding residual models computed by the difference between the elastic starting models (Figure 6.4(c,g)) and the final FWI models. The perturbations found with Test *H* (Figure 6.17d) are laterally continuous and present a lower wavenumber content than when the geophone data are inverted. To better visualize the structures elastic FWI creates, we show in Figure 6.18 some enlargements at different depths in the starting and the final models of Test *HG*. In addition to the Poisson's ratio we show the V_P/V_S and $V_P * V_S$ models (Figure 6.19). The V_P/V_S ratio allows one to discriminate areas

saturated in fluids and gives some insights into their physical state (in terms of pressure and temperature), while the $V_P * V_S$ quantity is representative of the lithological variations and to the different porosities and/or crack densities, minimizing at the same time the effects related to fluid saturation (Vanorio & Virieux, 2005).

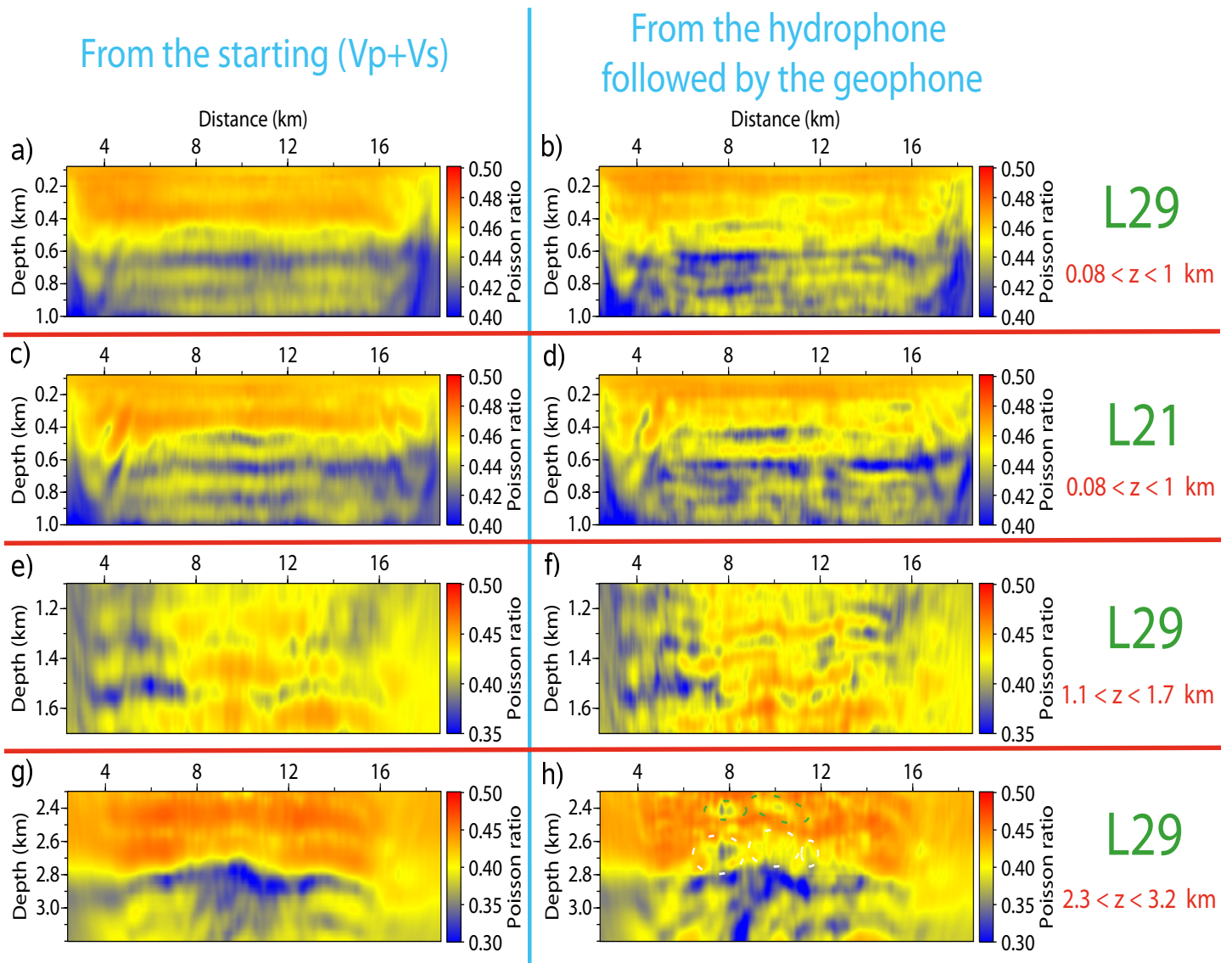


FIGURE 6.18 – Enlargements on the starting and final Poisson's models of Test *HG* for lines 21 and 29 at different range of depths. Note that the color scale can change from one layer to the next.

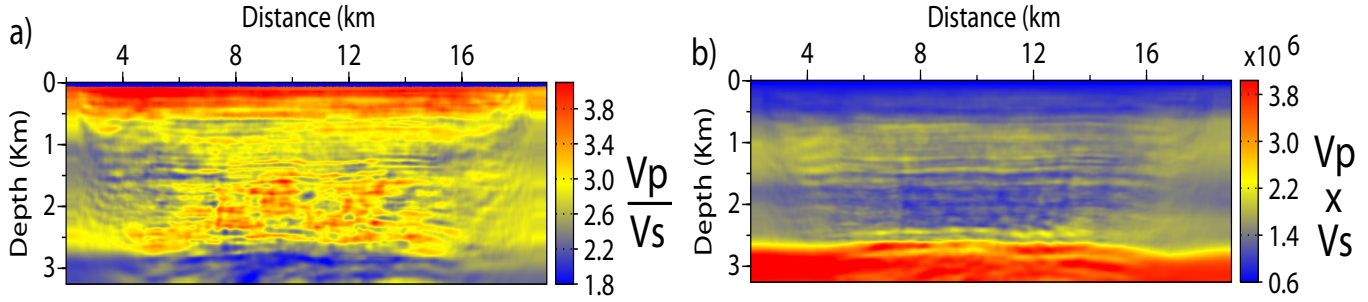


FIGURE 6.19 – Other combinations of P-wave velocity (V_P) and S-wave velocity (V_S) models. (a-b) V_P/V_S (a) and $V_P * V_S$ (b) inferred from the FWI models of Test *HG* for line 29.

6.9.2.1 Reflectors at 0.6 km, 1.25 km, 1.5 km and 2.4 km in depth

The $V_P * V_S$ values derived from the Test *HG* of line 29 (Figure 6.19b) are very weak and show little variations down to 0.6 km depth, while the corresponding V_P/V_S model (Figure 6.19a) is more contrasted at these depths. This supports that the first 0.6 km in depth correspond to soft unconsolidated quaternary sediments presenting a high homogeneous porosity, while the V_P/V_S variations are related to differences in fluid saturation. The Poisson model for the Test *HG* (Figures 6.17e and 6.18(a-f)) and the V_P/V_S map show some laterally continuous reflectors at 1.25 km and 1.5 km in depth, and some significant decrease of the Poisson's ratio at 0.6 km in depth at different locations of lines 21 and 29. A more detailed examination of the Poisson's ratio models suggests at this depth the presence of an upper layer with a high Poisson ratio value above a second layer with a smaller Poisson's ratio. This supports our previous interpretation that gas (in the lower layer) accumulates below lithological barriers. The same interpretation can be proposed for the presence of two strong (almost continuous) positive and negative perturbations of the Poisson's ratio at 2.4 km in depth at $X = 8$ km and $X = 10.5$ km of distance along line 29 (Figures 6.17f and 6.18h, green ellipses). It is worth noting a good correlation between the Poisson's ratio structure and the Q_P and V_P FWI models in which are found lower values at the same location (Figure 6.20f).

6.9.2.2 Reservoir level

The reservoir area is of particular interest as the elastic FWI generates the strongest perturbations at this level (Figure 6.17f). The $V_P * V_S$ map of line 29 (Figure 6.19b) shows the stack of an upper layer with a high $V_P * V_S$ value overlying a thin layer with a smaller $V_P * V_S$ value between 2.6 km and 2.8 km in depth. As $V_P * V_S$ is sensitive to the porosity, we propose that the upper and lower layers correspond respectively to the cap rock of the reservoir and the reservoir, where higher porosities are expected. In the enlargement of Figure 6.18h (white ellipse), we highlight between 2.6 and 2.7 km in depth some vertical features with Poisson ratio values lower than in the starting model (Figure 6.18g). These features might be interpreted as gas-filled fractures skewering the cap rock of the anticlinal, as we know that the gas in the upper tertiary sediments originates from the reservoir level because the cap rock is not fully hermetic (Munns, 1985). Of note, the in-line position of these features match the location where we have shown some lower values of V_P , ν and Q_P at 2.4 km in depth. Some pinpoints areas

with very low value of the Poisson ratio are also shown below 2.7 km in depth. These features can be firstly related to the very high level of fluid saturation in the reservoir, reaching up to 95 % (Munns, 1985). This high saturation is permitted by an important porosity, in some parts higher than 50 %. Furthermore, the reservoir area is known to be partitioned by faults created by the successive tectonic phases, and the porosity can vary significantly from a compartment to the next. Analysis of crossed dipole sonic logs in the chalk section by Mueller *et al.* (1997) have shown that V_S is strongly affected by the porosity, with up to 35 % reduction in velocity as porosity increase by 30 %. Vertical cracks and fractures are other possible causes of the Poisson's ratio decrease, independently to the fluids saturation (Fortin *et al.*, 2007). Azimuthal anisotropy is an indicator of porosity change and fracturing. Mueller *et al.* (1997) find from sonic logs that azimuthal anisotropy is of the order of 5-10% in magnitude in the upper chalk, where higher porosity is expected. Azimuthal anisotropy was also measured from shear wave splitting by Olofsson *et al.* (2003) and has been used for 4D-AVOA studies by Barkved *et al.* (2003). Because the V_P/V_S ratio shows more contrasts in the reservoir than the $V_P * V_S$ map, we propose that fluid saturation is more significant than change in porosity in the reservoir.

6.9.3 I_P, I_S, Q_S parameters

The impedance models inferred from the corresponding velocity models and density for line 29 clearly show the superimposition of the cap rock of the reservoir characterized by a high impedance over the reservoir characterized by a lower impedance between 2.6 km and 2.8 km depth. This is consistent with the study of Barkved *et al.* (1999), who propose that the strong porosity of the reservoir is responsible for a low value of I_P in the chalk, which can reach a value equal or even smaller than that found in the tertiary sediments. We compute the values of the Q_S parameter for lines 21 and 29 (Figure 6.20(d-e)) using the empirical law of Waters (1978) given by $Q_S = (4/3)Q_P(V_S/V_P)^2$. This expression relies on the assumption that there is no dissipation during a purely compressional cycle. For seismic frequencies, P attenuation is related to the "squirt" of fluid within the pore space, especially if the pore fluid is compressible (Thomsen *et al.*, 1997). Pore fluid compressibility is of second-order importance for V_S , suggesting that Q_S is not supposed to decrease within gas charged sediments. This is mostly the case between 1.5 km and 2.6 km depth for Q_S derived from line 21, but we observe rather low values for Q_S from line 29 at these depths. These models could nevertheless be used as realistic starting models in a further FWI involving Q_S .

6.9.4 Hydrocarbon indicators

We now derive two popular hydrocarbon indicators known as the difference between the P and S-wave reflection coefficients ($R_P - R_S$), and the AVO product (Castagna & Smith, 1994). The reflection coefficients are given by the formula :

$$R_P = \left(\frac{\Delta V_P}{V_P} + \frac{\Delta \rho}{\rho} \right) / 2 \quad ; \quad R_S = \left(\frac{\Delta V_S}{V_S} + \frac{\Delta \rho}{\rho} \right) / 2, \quad (6.8)$$

where

$$\begin{aligned} V_P &= (V_{P,2} + V_{P,1})/2 \quad ; \quad \Delta V_P = V_{P,2} - V_{P,1}, \\ V_S &= (V_{S,2} + V_{S,1})/2 \quad ; \quad \Delta V_S = V_{S,2} - V_{S,1}, \\ \rho &= (\rho_2 + \rho_1)/2 \quad ; \quad \Delta \rho = \rho_2 - \rho_1. \end{aligned} \quad (6.9)$$

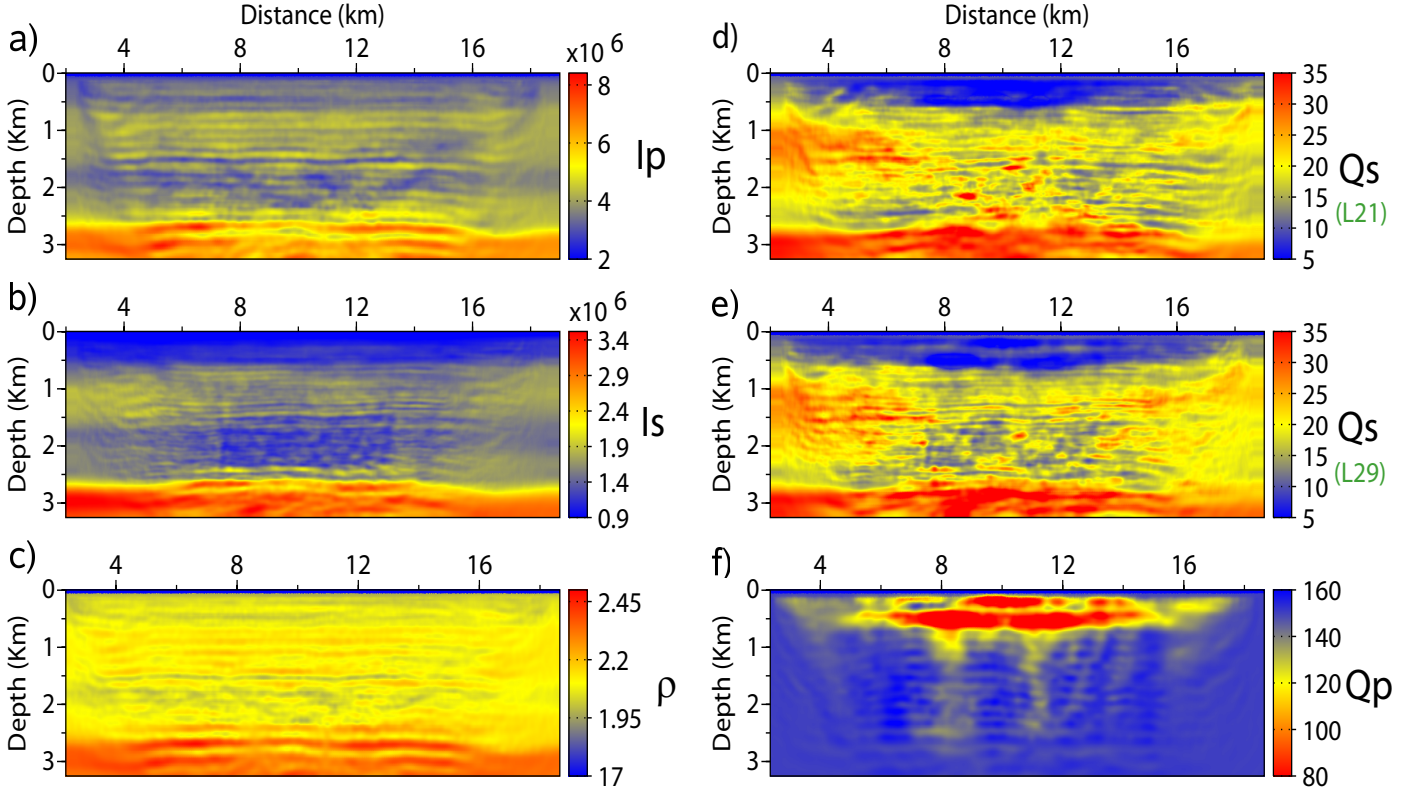


FIGURE 6.20 – (a-b) Compressive and shear impedance models of line 29; (d,e) Q_s model of line 21 and 29 computed using the formula from Waters (1978) ($Q_s = (4/3)Q_p (V_s/V_p)^2$). The impedances and the Q_s parameters are derived using the final V_p and V_s FWI models inferred from Test *HG*, as well as the density and attenuation models shown in (c,f) previously derived by acoustic FWI.

The difference between the reflection coefficients can thus be expressed as $R_P - R_S = (\Delta V_P/V_P - \Delta V_S/V_S)/2$. The AVO product (AVO_P) is defined as

$$AVO_P = R_P \left[-2 \left(\frac{V_S}{V_P} \right)^2 \frac{\Delta \rho}{\rho} + \frac{1}{2} \frac{\Delta V_P}{V_P} - 4 \left(\frac{V_S}{V_P} \right)^2 \frac{\Delta V_S}{V_S} \right]. \quad (6.10)$$

This formula allowed Djikpéssé & Tarantola (1999b) to clearly image a shale over gas sand reservoir. We show in Figure 6.21 these two quantities derived from the starting models used for the acoustic and elastic FWI, and from the FWI models of Test *HG* for the lines 21 and 29. The $R_P - R_S$ images clearly show the reflectors where a more important accumulation of gas occurs (Figure 6.21(a,c,e,g), white reflectors). The AVO product allows to highlight the layers with an hydrocarbon potential (Figure 6.21(b,d,f,h), white reflectors). The reservoir level consequently shows the strongest amplitude, although we note the acoustic starting model already present a white pattern in this area. In the final AVO product map of line 21, it is interesting to note the presence of lower values bending on the bottom right at the position of $X = 12$ km at 1 km in depth (see the white dashed circles) (Figure 6.21h). This feature is consistent with the low Q_P area that we previously highlighted in the Paper I and that we interpreted as a gas-filled fracture.

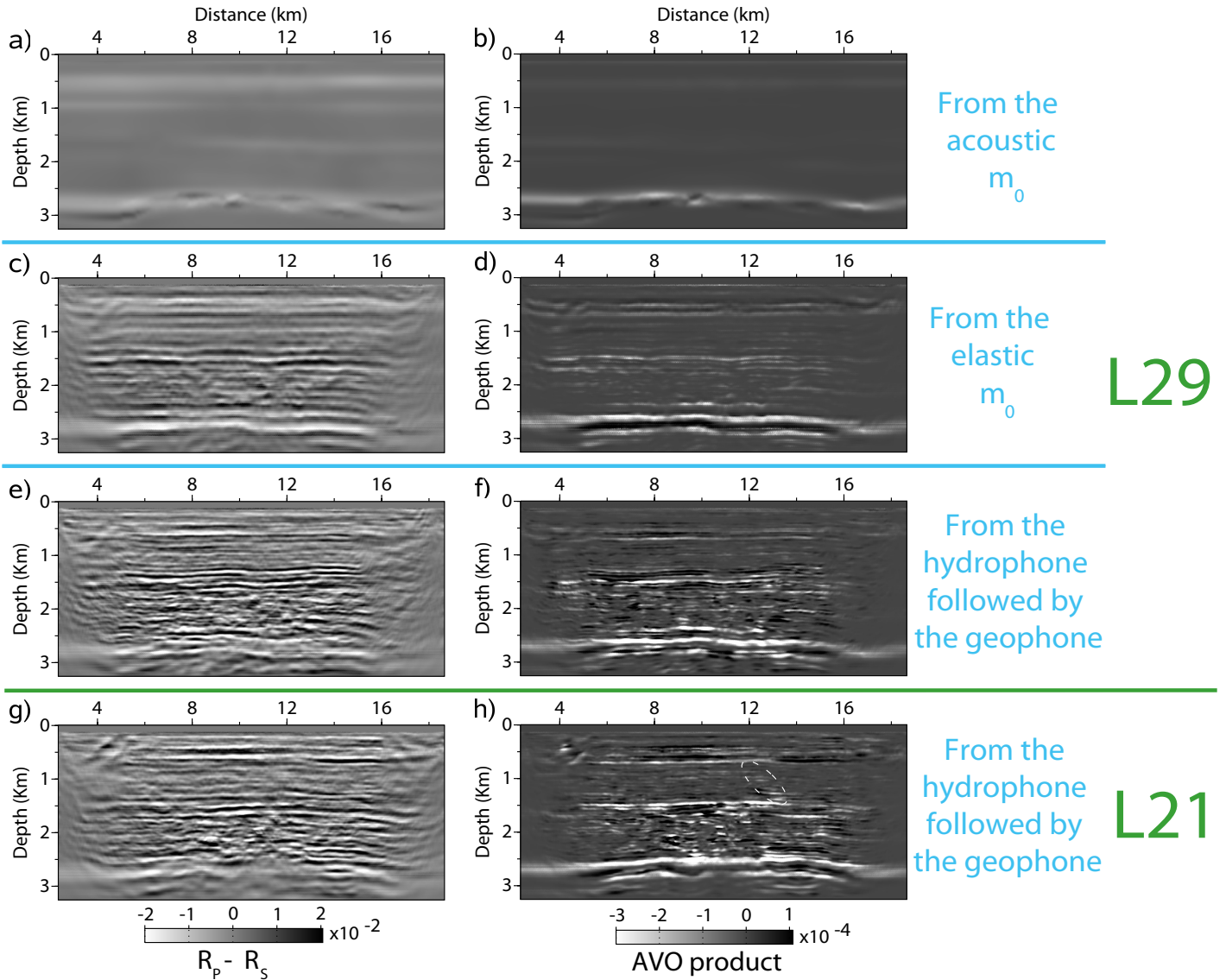


FIGURE 6.21 – Difference between the P and S reflection coefficients on the left panel, and AVO product on the right panel (see definition in the text). The two quantities are computed from the acoustic (a-b) and elastic (c-d) starting models, and from the final models of Test *HG* for line 29 (e-f) and line 21 (g-h).

6.10 Conclusion

We show the feasibility of visco-elastic VTI FWI of multi-component OBC data recorded in the Valhall oil field for the reconstruction of V_{P0} , V_S , ρ and Q_P . We have proposed a FWI workflow, which proceeds hierarchically over parameter classes and data components, to reduce the non linearity of the multiparameter inversion. All the criteria to appraise the results highlighted the benefit to invert V_{P0} and V_S from the hydrophone before the geophone data, to recover the intermediate wavenumbers before the higher ones. Compared to mono-

parameter acoustic imaging, the reconstruction of visco-elastic parameters provide invaluable inferences on lithological barriers below which gas accumulates, on gas injected fractures and on porosity and fluid saturation. A perspective of work is to update in a ultimate step the previously inverted parameters together with Q_S , within the visco-elastic approximation as we know it better predicts the AVO behavior than the visco-acoustic approximation. It would be also interesting to jointly invert the velocity and the pressure components (with 8 s of signal) in this ultimate step.

Acknowledgments

This study was funded by the SEISCOPE consortium <http://seiscope.oca.eu>, sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The linear systems were solved with the MUMPS package, available on <http://graal.ens-lyon.fr/MUMPS/index.html> and <http://mumps.enseeiht.fr>. The mesh generation was performed with help of TRIANGLE, available on <http://www.cs.cmu.edu/~quake/triangle.html>. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and of CINES/IDRIS under the allocation 2010- [project gao2280] made by GENCI (Grand Equipement National de Calcul Intensif), and we gratefully acknowledge both of these facilities and the support of their staff. We also thank BP Norge AS and their Valhall partner Hess Norge AS for allowing access to data to the Valhall data set as well as the well-log velocities.

6.11 Conclusion aux chapitres multiparamètres

L'inversion des formes d'ondes multiparamètre est un sujet stimulant aux enjeux majeurs. Nous avons illustré à travers nos applications sur cas synthétique et réel ses principales difficultés, liées à la sensibilité différente des données aux paramètres et aux unités des paramètres. Afin de palier à ces problèmes, nous avons promu une méthode d'optimisation basée sur une régularisation de type Tikhonov, et discuté des différentes stratégies d'inversion. Nous avons montré la faisabilité de l'inversion des paramètres V_P , V_S , ρ et Q_P aux données réelles, et avons utilisé différents critères de qualité pour évaluer nos résultats. Les interprétations géologiques et les déductions pétrophysiques que nous avons pu en faire illustrent les enjeux majeurs de l'inversion multiparamètre en terme de reconnaissance et de qualification des ressources de la subsurface.

Chapitre 7

Conclusions et perspectives

Mon travail de thèse a été réalisé dans le cadre du consortium SEISCOPE consacré à l'imagerie sismique haute résolution en prenant en compte tout le signal sismique. Cette approche intitulée inversion des formes d'onde complète a été explorée par divers chercheurs et étudiants de ce consortium et mon travail s'inscrit dans cette dynamique d'exploration des potentialités de cette méthode. Au terme de ce travail de thèse, j'aimerais conclure sur les points qui me paraissent importants et sur les perspectives qu'il est possible de dessiner dès à présent.

7.1 Un travail s'inscrivant dans la prise en compte des dispositifs grands-angles

Mon travail de thèse, à travers les différents thèmes abordés, s'est inscrit dans la tendance initiée par G. Pratt, faisant la promotion des dispositifs grands-angles. Ce thème central s'est progressivement dégagé au cours de ma thèse, notamment grâce à l'application de l'inversion des formes d'ondes et de la stéréotomographie aux données du champ de Valhall, présentant un déport maximum de 13 km. Tout d'abord, je montre l'incapacité des modèles isotropes à prédire conjointement la cinématique des arrivées à courts et grands offsets, quand la propagation est celle effectuée dans un milieu anisotrope VTI. J'illustre ensuite sur un cas synthétique que la prise en compte de très grands offsets (32 km) permet d'améliorer l'illumination du milieu en profondeur. Dans le même temps, ces données peuvent aussi être à l'origine de non-linéarités se traduisant par l'apparition d'artefacts dans le modèle reconstruit, car l'onde interagit fortement avec les hétérogénéités du milieu (Sirgue, 2006). Ces artefacts peuvent être réduits en utilisant des approches hiérarchiques permettant d'introduire progressivement dans l'inversion des sous-ensembles de données au pouvoir de résolution croissant (Sirgue, 2003; Brenders & Pratt, 2007b; Brossier *et al.*, 2009a; Shin & Cha, 2009). De la même manière, je montre lors des applications stéréotomographiques qu'une stratégie basée sur l'introduction séquentielle des données stéréotomographiques associées à des angles d'ouverture de plus en plus faibles, permet de réduire la non-linéarité de l'inversion conjointe des données courts et grands angles. Par ailleurs, j'illustre comment les imprécisions des paramètres anisotropes et d'atténuation se traduisent dans les sismogrammes par de plus forts résidus à grands offsets. Enfin, je montre lors de l'inversion élastique des données multicomposantes l'apport des données grands-angles pour la reconstruction des longueurs d'onde intermédiaires des vitesses de cisaillement, à partir de l'AVO des réflexions PP qui sont sensibles à la variation de la vitesse des ondes S.

7.2 Implications de mon travail

Dans un premier temps, je montre l'importance de la prise en compte de l'anisotropie en inversion des formes d'ondes, pour un meilleur positionnement des réflecteurs en profondeur. Sur le champ de Valhall, un écart de profondeur de près de 400 m est observé au niveau du réflecteur dit *profond* (à 3.5 km de profondeur), entre les images migrées isotrope et anisotrope. Cette prise en compte est donc indispensable dans les problématiques de forages. Je mets aussi l'accent sur le fait que le sens donné à une vitesse obtenue par un processus d'imagerie dépend directement de celui-ci (quelle approximation ?) et du dispositif d'acquisition utilisé (court- ou grand-angle ?). Je montre ainsi que la vitesse reconstruite par un processus d'imagerie isotrope dans un milieu anisotrope peut avoir un sens différent près de la surface et en profondeur, où elle est respectivement représentative des vitesses horizontales et NMO. Ceci est une conséquence directe de la variation de l'éclairage angulaire du milieu avec la profondeur, cet éclairage devenant de plus en plus étroit vers les faibles ouvertures au fur et à mesure que la profondeur augmente. Ainsi, l'imagerie est majoritairement dominée par l'influence des arrivées grands-angles dans la partie superficielle, tandis que la partie profonde est contrainte par les ondes réfléchies de faibles ouvertures angulaires.

Il est clairement apparu que la stéréotomographie s'avère être un outil de choix pour la construction d'un macromodèle de vitesse initial à l'inversion des formes d'ondes, d'une part car la stratégie proposée (sur les angles d'ouverture) permet de combiner les arrivées réfléchies et réfractées, dont la cinématique est d'un grand intérêt pour l'inversion des formes d'ondes et d'autre part, parce que cette méthode tire bénéfice d'un pointé semi-automatique sur des données hautes fréquences, permettant de pointer un grand nombre d'arrivées et donc d'illuminer de manière complète le milieu. La stratégie proposée permet en outre d'améliorer les résultats obtenus jusqu'alors à partir des données à courts offsets.

Par ailleurs, je montre l'intérêt des imageries multiparamètres pour la caractérisation des fluides avec une étude de faisabilité de la reconstruction des vitesses P et S, de la densité et de l'atténuation sur les données du champ de Valhall. Je propose une méthodologie fondée sur une reconstruction hiérarchique des classes de paramètres et sur l'introduction progressive des différentes composantes des données dans l'inversion. L'approche hiérarchique sur les classes de paramètres consiste à mettre à jour dans un premier temps les paramètres ayant une empreinte dominante dans les données avant la reconstruction des paramètres de second ordre. En pratique sur Valhall, j'inverse V_P avant la densité et l'atténuation, et ces trois paramètres acoustiques avant l'inversion de la vitesse V_S . Je montre ensuite l'intérêt de reconstruire hiérarchiquement V_S à partir des données des hydrophones puis des géophones. Je préconise aussi d'utiliser une paramétrisation définie par les paramètres $(V_{P0}, V_S, \epsilon, \delta, \rho, Q_P, Q_S)$, garantissant une reconstruction haute résolution des vitesses. Je justifie mon choix de paramétrisation en vitesse-densité plutôt qu'en vitesse-impédance sur un cas synthétique, où je montre que cette dernière paramétrisation ne permet pas une reconstruction haute résolution de la vitesse, et induit des instabilités dans le modèle de densité déduit des paramètres de vitesse et d'impédance inversés. Suite au travail de Brossier (2009), je recommande la méthode d'optimisation L-BFGS, en particulier dans le cadre des inversions multiparamètres. Cette approche fournit en effet une estimation économique du Hessien permettant de mettre à l'échelle les gradients associés à chaque classe de paramètres, de rendre compte des couplages entre les paramètres et des effets de diffractions doubles. Je montre que le L-BFGS est instable sur les données réelles s'il n'est pas couplé à une régularisation de type Tikhonov, comportant des termes de contraintes et

de lissage dans l'espace des modèles. Cette régularisation s'avère aussi utile pour rendre compte de la sensibilité variable des données aux différentes classes de paramètres secondaires, grâce au choix adapté des paramètres d'amortissements de l'inversion. Ce choix est réalisé par essai erreur de sorte que le résultat satisfasse à notre a priori géologique. Afin d'illustrer le pouvoir de l'imagerie multiparamètre à renseigner sur la pétrophysique et la géologie du milieu, je déduis diverses quantités physiques issues des paramètres inversés, et je propose des interprétations géologiques cohérentes avec l'histoire et nos connaissances du champ de Valhall.

7.3 Perspectives

La reconstruction des paramètres anisotropes par inversion des formes d'ondes est un sujet novateur et très difficile (Gholami *et al.*, 2010a,b; Plessix & Cao, 2011), les données n'étant sensibles qu'aux grandes longueurs d'ondes de l'anisotropie et très peu sensibles aux courtes longueurs d'ondes. La non-linéarité d'une inversion anisotrope est illustrée dans notre étude : un grand nombre de modèles reconstruits peuvent expliquer à l'identique les données anisotropes. De plus amples études sont nécessaires pour trouver les paramétrisations et les paramètres à inverser les plus adaptés, notamment dans l'approximation élastique anisotrope, ainsi que sur les stratégies pour intégrer des informations a priori tel que le calage aux puits des réflecteurs, s'avérant indispensable pour mieux contraindre ces paramètres.

Nous avons pu constater les limites de l'outil isotrope de stéréotomographie à notre disposition lors de son application aux données anisotropes du champ de Valhall. Il serait intéressant d'étendre l'outil aux cas des milieux anisotropes (Nag *et al.*, 2006, 2010) et d'appliquer notre stratégie aux mêmes données, pour comparer sa capacité de résolution relativement au modèle construit par tomographie en réflexion par BP. Il pourrait aussi être envisagé de l'étendre en 3 dimensions (Chalard *et al.*, 2002) et à l'utilisation de données converties PS (Alerini *et al.*, 2002). Il faudrait cependant pour appliquer cet outil au champ de Valhall des données prétraitées sans multiples. Une autre extension riche en applications, notamment dans la communauté académique ou pour compenser une acquisition moins dense suivant la direction *cross-line* des acquisitions 3D, serait de pouvoir prendre en compte des acquisitions plus lâches (Alerini *et al.*, 2009).

D'autres études de sensibilité pourraient être envisagées sur le volet multiparamètre du champ de Valhall. Je pense notamment à l'application d'amortissements temporels des données pour réduire la non-linéarité de l'inversion, ou l'application d'un opérateur de pondération des vecteurs résidus dans la fonction coût tels que des gains en offset (Brossier *et al.*, 2009c), pour mieux extraire l'information des données converties PS. Il serait aussi intéressant d'inverser des données de géophones comportant plus de 8 s de signal, sur lesquelles serait enregistrée l'arrivée SS s'étant réfléchi au toit du réservoir. Une autre perspective serait d'inverser avec l'approximation élastique le paramètre Q_S conjointement à l'ensemble des paramètres déjà inversés. Le modèle initial de Q_S pouvant être obtenu par la loi de Waters (1978) (Figure 6.20). Comme l'approximation élastique modélise de manière plus appropriée le comportement AVO des ondes, la reconstruction de la densité et de l'atténuation pourrait s'en trouver améliorée. Par ailleurs, Jin *et al.* (2000) et Djikpéssé & Tarantola (1999b) suggèrent d'inverser la densité à partir de données converties PS plutôt qu'à partir des ondes P, car la densité serait plus sensible à ce type de données. D'autre part, Sears *et al.* (2008) montre que la reconstruction de V_S à partir des réflexions PP est plus stable lorsqu'elle est effectuée à partir de la composante

verticale du géophone (V_Z) plutôt qu'à partir du signal enregistré par l'hydrophone, car V_Z serait moins sensible aux multiples.

L'inversion des formes d'ondes en trois dimensions s'est révélée ces dernières années un outil très puissant pour réduire les non-linéarités de la méthode et obtenir des images hautes résolutions. On peut se demander dans quelle mesure la troisième dimension améliorera les résultats multiparamètres obtenus en deux dimensions.

L'inversion des formes d'onde complètes est censée imager la réflectivité si l'inversion peut être poussée à des fréquences suffisamment élevées. Cependant, les applications de la méthode aux données réelles ne sont jusqu'à maintenant appliquées qu'à des données relativement basses fréquences. Cela est en partie lié au coût numérique plus élevé de la modélisation aux fréquences plus hautes, mais probablement aussi à cause de la limitation de la validité de l'approche linéarisée aux hautes fréquences, les ondes étant alors plus sensibles aux diffractions doubles. Un test préliminaire d'inversion de V_P , la densité et l'atténuation aux fréquences supérieures à 7 Hz (en approximation acoustique) sur le champ de Valhall, a montré que l'inversion n'a que légèrement perturbé l'atténuation en surface, et a laissé les autres paramètres inchangés. Cette question reste donc à approfondir à l'avenir, en réitérant par exemple ce test avec l'approximation élastique.

Différentes pistes sont actuellement explorées pour réduire la non-linéarité de la méthode d'inversion des formes d'onde complètes. À cet effet, des alternatives à la régularisation Tikhonov, qui a pour but de trouver des images lisses du milieu, peuvent être employées. La régularisation de type *Total Variation* (Akçelik, 2002; Hu *et al.*, 2009) ou la régularisation multiplicative (van den Berg *et al.*, 2003; Abubakar *et al.*, 2009) visent au contraire à accentuer les contrastes de l'image.

La prise en compte des informations a priori est un autre sujet très intéressant (Tonellot & Broadhead, 2011). Des opérateurs de lissage directionnels (Hale, 2011) le long des interfaces géologiques peuvent être employés pour mieux mettre en évidence les structures, tout en préservant les discontinuités telles que des failles. Les interfaces géologiques sont connues a priori par des logs de puits ou par des images migrées. (Guitton *et al.*, 2010, 2012) intègre ce type d'opérateurs dans le formalisme de l'inversion des formes d'onde complètes. Il montre son utilité en présence d'un jeu de données décimées et sur un modèle présentant un fort contraste de vitesses (un dôme de sel) : la reconstruction s'avère pertinente en présence de données bruitées, irrégulières et présentant un déficit en basses fréquences.

Des alternatives aux fonctions coûts traditionnelles ont aussi été proposées, basées sur une mesure de corrélation (Leeuwen & Mulder, 2010) ou sur les attributs d'enveloppe (Fichtner *et al.*, 2008; Bozdag *et al.*, 2011). La fonction coût standard attache en effet un poids plus important à la partie du signal ayant une forte amplitude, l'avantage étant qu'elle est moins sensible au bruit. Une fonctionnelle basée sur la notion d'enveloppe et de phase instantanée (Bozdag *et al.*, 2011) permet de séparer l'amplitude et la phase, pour extraire de l'information sur des phases de faibles amplitudes. Cette approche s'avère cependant sensible au bruit et aux sauts de phase. Par ailleurs, nous avons vu dans le chapitre 3 que la répétabilité des ondelettes sources estimées par point de tir est un indicateur de la qualité du modèle (Malinowski *et al.*, 2011). Ce critère pourrait aussi être utilisé dans la fonction coût.

L'inversion des formes d'ondes est une approche basée sur l'optimisation d'une fonctionnelle dans l'espace des données, alors que la fonction coût des méthodes d'analyse de vitesse dirigée par la migration opère dans l'espace migré pour aplatir les CIGs (Symes, 2008). Une inversion

des formes d'onde complètes dans l'espace des modèles a été proposée pour combiner ces deux points de vue. Nous pouvons citer la méthode dite de *Shot-Based Differential Semblance Optimization* (Symes & Carazzone, 1991; Symes, 1993; Plessix, 2006) ou celle des CIG étendus en temps ou en espace (Yang & Sava, 2011a,b).

Ainsi, la méthode d'inversion des formes d'onde complètes fait l'objet d'intenses développements et a encore un bel avenir devant elle, le champ de recherche étant encore très vaste et plein de promesses. J'espère que mon travail en est une belle illustration.

Bibliographie

- Abubakar, A., Hu, W., Habashy, T. M. & van den Berg, P. M. (2009). Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data. *Geophysics*, 74(6):WCC47–WCC58.
- Ainsworth, M., Monk, P. & Muniz, W. (2006). Dispersive and Dissipative Properties of Discontinuous Galerkin Finite Element Methods for the Second-Order Wave Equation. *Journal of Scientific Computing*, 27(1-3):5–40.
- Akçelik, V. (2002). *Multiscale Newton-Krylov methods for inverse acoustic wave propagation*. Thèse de doctorat, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- Aki, K. & Richards, P. (1980). *Quantitative Seismology : Theory and Methods*. W. H. Freeman & Co, San Francisco.
- Aki, K. & Richards, P. G. (2002). *Quantitative seismology, theory and methods, second edition*. University Science Books, Sausalito, California.
- Alerini, M., Bégat, S. L., Lambaré, G. & Baina, R. (2002). 2d pp- and ps-stereotomography for a multicomponent dataset. *SEG Technical Program Expanded Abstracts*, 21(1):838–841.
- Alerini, M., Lambaré, G., Baina, R., Podvin, P. & Bégat, S. L. (2007). Two-dimensional PP/PS-stereotomography : P- and S-waves velocities estimation from OBC data. *Geophysical Journal International*, 170:725–736.
- Alerini, M., Traub, B., Ravaut, C. & Duveneck, E. (2009). Prestack depth imaging of ocean-bottom node data. *Geophysics*, 74(6):WCA57–WCA63.
- Alkhalifah, T. & Tsvankin, I. (1995). Velocity analysis for transversely isotropic media. *Geophysics*, 60:1550–1566.
- Amestoy, P., Davis, T. A. & Duff, I. S. (1996). An approximate minimum degree ordering algorithm. *SIAM Journal of Matrix Analysis and Applications*, 17:886–905.
- Amestoy, P. R., Duff, I. S. & L’Excellent, J. Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering*, 184:501–520.
- Artman, B. (2007). *Passive Seismic Imaging*. Thèse de doctorat, Stanford University, Stanford Exploration Project report SEP 128.

- Askan, A., Akcelik, V., Bielak, J. & Ghattas, O. (2007). Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures. *Bulletin of the Seismological Society of America*, 97(6):1990–2008.
- Backus, G. E. (1962). Long-wave elastic anisotropy produced by horizontal layering. *Journal Geophysical Research*, 67:4427–4440.
- Barkved, O., Albertin, U., Heavey, P., Kommedal, J., van Gestel, J., Synnove, R., Pettersen, H. & Kent, C. (2010a). Business impact of full waveform inversion at Valhall. In *Expanded Abstracts, 91 Annual SEG Meeting and Exposition (October 17-22, Denver)*, pages 925–929. Society of Exploration Geophysics.
- Barkved, O., Buer, K., Halleland, K., Kjelstadli, R., Kleppan, T. & Kristiansen, T. (2003). 4d seismic response or primary production and waste injection at the valhall field. In *Extended Abstracts, 65th Annual EAGE Conference & Exhibition*.
- Barkved, O., Heavey, P., Kommedal, J. H., van Gestel, J.-P., ve, R. S., Pettersen, H., Kent, C. & Albertin, U. (2010b). Business impact of full waveform inversion at valhall. *SEG Technical Program Expanded Abstracts*, 29(1):925–929.
- Barkved, O., Mueller, M. & Thomsen, L. (1999). Vector interpretation of the valhall 3d/4c obs dataset. In *Extended Abstracts, 61th Annual EAGE Conference & Exhibition, Helsinki, Finland*.
- Barkved, O., van Gestel, J., Bergsvik, L., Stockden, I. & Kommedal, J. (2009). Seismic plt - linking seismic time-lapse responses to production and injection data. In *Expanded Abstracts*.
- Barnes, C. & Charara, M. (2009). The domain of applicability of acoustic full-waveform inversion for marine seismic data. *Geophysics*, 74(6):WCC91–WCC103.
- Baysal, E., Kosloff, D. & Sherwood, J. (1983). Reverse time migration. *Geophysics*, 48:1514–1524.
- Ben-Hadj-Ali, H., Operto, S. & Virieux, J. (2008). Velocity model building by 3d frequency-domain, full-waveform inversion of wide-aperture seismic data. *Geophysics*, 73:WE101–WE117.
- Ben Jemaa, M., Glinsky-Olivier, N., Cruz-Atienza, V. M. & Virieux, J. (2009). 3D Dynamic rupture simulations by a finite volume method. *Geophysical Journal International*, 178:541–560.
- Berenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.
- Berkhout, A. J. & Wapenaar, C. P. A. (1989). One-way versions of the kirchhoff integral. *Geophysics*, 54:460–467.
- Bernacki, M. & Piperno, S. (2003). Méthode de type galerkin-discontinu pour la propagation des ondes en aéroacoustique. Rapport de Recherche 4932, INRIA Sophia-Antipolis.
- Beylkin, G. (1985). Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *Journal of Mathematical Physics*, 26:99–108.

- Bichler, A., Bobrowsky, P., Best, M., Douma, M., Hunter, J., Calvert, T. & Burns, R. (2004). Three-dimensional mapping of a landslide using a multi-geophysical approach : the quesnel forks landslide. *Landslides*, 1:29–40.
- Billette, F. & Lambaré, G. (1998). Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, 135(2):671–680.
- Billette, F., Le Bégat, S., Podvin, P. & Lambaré, G. (2003). Practical aspects and applications of 2D stereotomography. *Geophysics*, 68:1008–1021.
- Biondi, B. L. (2006). *3D seismic imaging*. SEG.
- Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T. & Spinder, D. A. (1985). Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50:903–923.
- Bleibinhaus, F., Hole, J. A., Ryberg, T. & Fuis, G. S. (2007). Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging . *Journal of Geophysical Research*, 112(B06315):doi :10.1029/2006JB004611.
- Bleistein, N. (1986). Two and one half dimensional in plane wave propagation. *Geophysical Prospecting*, 34:686–703.
- Bleistein, N. (1987). On the imaging of reflectors in the Earth. *Geophysics*, 52(7):931–942.
- Bohlen, T. & Saenger, E. H. (2006). Accuracy of heterogeneous staggered-grid finite-difference modeling of Rayleigh waves. *Geophysics*, 71:109–115.
- Bouchon, M., Campillo, M. & Gaffet, S. (1989). A boundary integral equation - discrete wavenumber representation method to study wave propagation in multilayered media having irregular interfaces. *Geophysics*, 54:1134–1140.
- Bozdag, E., Trampert, J. & Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2):845–870.
- Brenders, A. J. & Pratt, R. G. (2007a). Efficient waveform tomography for lithospheric imaging : implications for realistic 2D acquisition geometries and low frequency data. *Geophysical Journal International*, 168:152–170.
- Brenders, A. J. & Pratt, R. G. (2007b). Full waveform tomography for lithospheric imaging : results from a blind test in a realistic crustal model. *Geophysical Journal International*, 168:133–151.
- Brossier, R. (2009). *Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d'onde : développements méthodologiques et applications*. Thèse de doctorat, Université de Nice-Sophia-Antipolis.
- Brossier, R. (2011a). Two-dimensional frequency-domain visco-elastic full waveform inversion : Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444 – 455.

- Brossier, R. (2011b). Two-dimensional frequency-domain visco-elastic full waveform inversion : Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444 – 455.
- Brossier, R., Etienne, V., Operto, S. & Virieux, J. (2010a). Frequency-domain numerical modelling of visco-acoustic waves based on finite-difference and finite-element discontinuous galerkin methods. In Dissanayake, D. W., éditeur : *Acoustic Waves*, pages 125–158. SCIYO.
- Brossier, R., Gholami, Y., Virieux, J. & Operto, S. (2010b). 2D frequency-domain seismic wave modeling in vti media based on a hp-adaptive discontinuous galerkin method. In *Expanded Abstracts, 72th Annual International Meeting, EAGE*.
- Brossier, R., Operto, S. & Virieux, J. (2009a). 2D elastic frequency-domain full-waveform inversion for imaging complex onshore structures. In *Expanded Abstracts*, page U019. EAGE.
- Brossier, R., Operto, S. & Virieux, J. (2009b). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):WCC63–WCC76.
- Brossier, R., Operto, S. & Virieux, J. (2009c). Two-dimensional seismic imaging of the Valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion. *SEG Technical Program Expanded Abstracts*, 28(1):2293–2297.
- Brossier, R., Operto, S. & Virieux, J. (2010c). Which data residual norm for robust elastic frequency-domain full waveform inversion ? *Geophysics*, 75(3):R37–R46.
- Brossier, R. & Roux, P. (2011). Seismic imaging by frequency-domain double-beamforming full-waveform inversion. In *Expanded Abstracts, 73th Annual EAGE Conference & Exhibition, Vienna*, page C039. EAGE.
- Brossier, R., Virieux, J. & Operto, S. (2008a). 2D frequency-domain elastic full-waveform inversion using a P_0 finite volume forward problem. In *Expanded Abstracts, 78th Annual SEG Conference & Exhibition, Las Vegas*. Society of Exploration Geophysics.
- Brossier, R., Virieux, J. & Operto, S. (2008b). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling. *Geophysical Journal International*, 175(2):541–559.
- Broto, K., Ehinger, A., Kommedal, J. & Folstad, P. (2003). Anisotropic travelttime tomography for depth consistent imaging of pp and ps data. *The Leading Edge*, pages 114–119.
- Bunks, C., Salek, F. M., Zaleski, S. & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- Byrd, R., Lu, P. & Nocedal, J. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16:1190–1208.
- Carcione, J. M. & Helle, H. (2004). The physics and simulation of wave propagation at the ocean bottom. *Geophysics*, 69(3):825–839.
- Cary, P. & Chapman, C. (1988). Automatic 1-D waveform inversion of marine seismic refraction data. *Geophysical Journal of the Royal Astronomical Society*, 93:527–546.
- Castagna, J. & Smith, S. (1994). Comparison of avo indicators : A modeling study. *Geophysics*, 59:1849–1855.

- Červený, V., Molotkov, I. A. & Pšenčík, I. (1977). *Ray Method in Seismology*. Charles University Press, Praha.
- Chalard, E., Podvin, P., Bégat, S. L., Berthet, P. & David, B. (2002). 3-D stereotomographic inversion on a real dataset. *In Expanded Abstracts*, pages 946–948.
- Chambers, K. & Kendall, J.-M. (2010). Investigation of induced microseismicity at Valhall using the Life of Field Seismic array. *The Leading Edge*, 29:290–295.
- Chan, W. & McMechan, G. (1994). 3d elastic prestack reverse time migration. *Geophysics*, 59:597–609.
- Chauris, H. (2000). *Analyse de vitesse par migration pour l'imagerie des structures complexes en sismique réflexion*. Thèse de doctorat, Ecole des Mines de Paris.
- Chauris, H. & Noble, M. (2001). 2D velocity macro model estimation from seismic reflection data by local Differential Semblance Optimization : applications on synthetic and real data. *Geophysical Journal International*, 144:14–26.
- Chauris, H., Noble, M. & Taillander, C. (2008). Which initial model do we need for full waveform inversion? *In Expanded Abstracts, 70th Annual EAGE Conference & Exhibition, WO11 : Full waveform inversion - Current status and perspectives, Roma*. EAGE.
- Chavent, G. (2009). *Nonlinear least squares for inverse problems*. Springer Dordrecht Heidelberg London New York.
- Choi, Y., Min, D. & Shin, C. (2008). Two-dimensional waveform inversion of multi-component data in acoustic-elastic coupled media. *Geophysical Prospecting*, 56(6):863–881.
- Claerbout, J. (1985). *Imaging the Earth's interior*. Blackwell Scientific Publication.
- Claerbout, J. F. (1970). Coarse grid calculations of wave in inhomogeneous media with application to delineation of complicated seismic structure. *Geophysics*, 35(3):407–418.
- Claerbout, J. F. (1976). *Fundamentals of Geophysical Data Processing*. McGraw-Hill Book Co.
- Cockburn, B., Karniadakis, G. & Shu, C.-W. (2000). The development of discontinuous galerkin methods. *In* B. Cockburn, G. K. & Shu, C.-W., éditeurs : *Discontinuous Galerkin methods theory ; computation and application*, Lecture notes in computational Science and Engineering, 11. Springer.
- Cockburn, B., Li, F. & Shu, C. W. (2004). Locally divergence-free discontinuous Galerkin methods for the Maxwell equations. *Journal of Computational Physics*, 194:588–610.
- Cohen, J. K. & Bleistein, N. (1979). Velocity inversion procedure for acoustic waves. *Geophysics*, 44:1077–1087.
- Cruse, E., Pica, A., Noble, M., McDonald, J. & Tarantola, A. (1990). Robust elastic non-linear waveform inversion : application to real data. *Geophysics*, 55:527–538.
- Cruse, E., Wideman, C., Noble, M. & Tarantola, A. (1992). Nonlinear elastic inversion of land seismic reflection data. *Journal of Geophysical Research*, 97:4685–4705.

- Dahlen, F. A., Hung, S. H. & Nolet, G. (2000a). Fréchet kernels for finite-difference traveltimes - I. theory. *Geophysical Journal International*, 141:157–174.
- Dahlen, F. A., Hung, S.-H. & Nolet, G. (2000b). Fréchet kernels for finite-frequency traveltimes - I. theory. *Geophysical Journal International*, 141:157–174.
- Dahlen, F. A. & Tromp, J. (1998). *Theoretical Global Seismology*. Princeton University Press, Princeton.
- Danecek, P. & Seriani, G. (2008). An efficient parallel Chebyshev pseudo-spectral method for large-scale 3D seismic forward modelling. *In Expanded Abstracts*, page P046.
- Dasgupta, R. & Clark, R. (1998). Estimation of q from surface seismic reflection data. *Geophysics*, 63:2120–2128.
- de Basabe, J., Sen, M. & Wheeler, M. (2008). The interior penalty discontinuous galerkin method for elastic wave propagation : grid dispersion. *Geophysical Journal International*, 175:83–93.
- De la Puente, J., Ampuero, J.-P. & Käser, M. (2009). Dynamic Rupture Modeling on Unstructured Meshes Using a Discontinuous Galerkin Method. *Journal of Geophysical Research*, 114:B10302.
- de la Puente, J., Käser, M., Dumbser, M. & Igel, H. (2007). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes IV : Anisotropy. *Geophysical Journal International*, 169(3):1210–1228.
- Debski, W. & Tarantola, A. (1995). Information on elastic parameters obtained from amplitude of reflected waves. *Geophysics*, 60:1426–1436.
- Dellinger, J. (2008). Low frequencies using conventional sensors : “sign-bit” recording revisited. volume 27, pages 149–153. SEG.
- Dellinger, J. & Yu, J. (2009). Low-frequency virtual point-source interferometry using conventional sensors. *In Expanded Abstracts*.
- Descloitres, M., Ruiz, L., Sekhar, M., Legchenko, A., Braun, J., Kumar, M. M. & Subramanian, S. (2008). Characterization of seasonal local recharge using electrical resistivity tomography and magnetic resonance sounding. *Hydrological Processes*, 22:384–394.
- Dessa, J. X. (2001). *Modélisation et inversion de la forme d’onde en domaine fréquentiel : applications à la tomographie multi-offsets*. Thèse de doctorat, Université de Paris Sud.
- Dessa, J. X., Operto, S., Kodaira, S., Nakanishi, A., Pascal, G., Uehira, K. & Kaneda, Y. (2004). Deep seismic imaging of the eastern Nankai trough (Japan) from multifold ocean bottom seismometer data by combined traveltime tomography and prestack depth migration. *Journal of Geophysical Research*, 109(B02111):doi :10.1029/2003JB002689.
- Devaney, A. J. (1982). A filtered backprojection algorithm for diffraction tomography. *Ultrasonic Imaging*, 4:336–350.
- Dix, C. H. (1955). Seismic velocities from surface measurements. *Geophysics*, 20:68–86.

- Djikipéssé, H. A. & Tarantola, A. (1999a). Multiparameter l_1 norm waveform fitting : Interpretation of gulf of mexico reflection seismograms. *Geophysics*, 64(4):1023–1035.
- Djikipéssé, H. A. & Tarantola, A. (1999b). Multiparameter l_1 norm waveform fitting : Interpretation of gulf of mexico reflection seismograms. *Geophysics*, 64(4):1023–1035.
- Domenico, S. N. (1984). Rock lithology and porosity determination from shear and compressional wave velocity. *Geophysics*, 49:1188–1195.
- Duff, I. S., Erisman, A. M. & Reid, J. K. (1986). *Direct methods for sparse matrices*. Clarendon Press, Oxford, U. K.
- Duff, I. S. & Reid, J. K. (1983). The multifrontal solution of indefinite sparse symmetric linear systems. *ACM Transactions on Mathematical Software*, 9:302–325.
- Dumbser, M. & Käser, M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes II : The Three-Dimensional Isotropic Case. *Geophysical Journal International*, 167(1):319–336.
- Ellefsen, K. J. (2009). A comparison of phase inversion and traveltimes tomography for processing near-surface refraction traveltimes. *Geophysics*, 74(6):WCB11–WCB24.
- Etienne, V. (2011). *Modélisation de la propagation d’ondes sismiques par éléments finis discontinus et inversion des formes d’onde dans les milieux élastiques à trois dimensions*. Thèse de doctorat, Université de Nice-Sophia-Antipolis, Sophia-Antipolis.
- Etienne, V., Chaljub, E., Virieux, J. & Glinsky, N. (2010a). An hp-adaptive discontinuous Galerkin finite-element method for 3D elastic wave modelling. *Geophysical Journal International*, 183(2):941–962.
- Etienne, V., Virieux, J., Hui, G., Jia, Y. & Operto, S. (2010b). Computational and methodological developments towards 3D full waveform inversion. In *EOS Trans. AGU, abstract S31A-2049*. American Geophysical Union, San Francisco, USA.
- Faccioli, E. F., Paolucci, R. & Quarteroni, A. (1997). 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. *Journal of Seismology*, 1:237–251.
- Farra, V. & Madariaga, R. (1988). Non-linear reflection tomography. *Geophysical Journal*, 95:135–147.
- Fichtner, A., Kennett, B. L. N., Igel, H. & Bunge, H. P. (2008). Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain. *Geophysical Journal International*, 175:665–685.
- Forgues, E. (1996). *Inversion linéarisée multi-paramètres via la théorie des rais*. Thèse de doctorat, Institut Français du Pétrole - University Paris VII.
- Forgues, E. & Lambaré, G. (1997). Parameterization study for acoustic and elastic ray+born inversion. *Journal of Seismic Exploration*, 6:253–278.
- Fornberg, B. (1996). *A Practical Guide to Pseudospectral Methods*. Cambridge Monographs on Applied and Computational Mathematics (No. 1).

- Fortin, J., Guéguen, Y. & Schubnel, A. (2007). Effects of pore collapse and grain crushing on ultrasonic velocities and vp/vs. *Journal of Geophysical Research*, 112.
- Freudenreich, Y., Singh, S. & Barton, P. (2001). Sub-basalt imaging using a full elastic wavefield inversion scheme. In *Extended Abstracts, 63th Annual EAGE Conference & Exhibition*.
- Futterman, W. (1962). Dispersive body waves. *Journal Geophysical Research*, 67:5279–5291.
- Gardner, G. H. F., Gardner, L. W. & Gregory, A. R. (1974). Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics*, 39:770–780.
- Garvin, W. W. (1956). Exact transient solution of the buried line source problem. *Proceedings of the Royal Society A*, 234:528–541.
- Gassmann, F. (1951). Über die elastizität poröser medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 96:1–23.
- Gauthier, O., Virieux, J. & Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms : numerical results. *Geophysics*, 51(7):1387–1403.
- Gélis, C. (2005). *Inversion de formes d'onde élastique dans le domaine espace-fréquence en deux dimensions. Application à la caractérisation de la subsurface dans le cadre de la détection de cavités souterraines*. Thèse de doctorat, Université Nice-Sophia Antipolis.
- George, A. & Liu, J. W. (1981). *Computer solution of large sparse positive definite systems*. Prentice-Hall, Inc.
- Gholami, Y., Brossier, R., Operto, S., Prioux, V., Ribodetti, A. & Virieux, J. (2011a). Two-dimensional acoustic anisotropic (VTI) full waveform inversion : the Valhall case study. *SEG Technical Program Expanded Abstracts*, 30:2543.
- Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. & Virieux, J. (2011b). Acoustic anisotropic full waveform inversion : sensitivity analysis and realistic synthetic examples. *SEG Technical Program Expanded Abstracts*, 30:2465.
- Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. & Virieux, J. (2012). Which parametrization for acoustic VTI full waveform inversion? - part 1 : sensitivity and trade-off analysis. *Geophysics*, Submitted.
- Gholami, Y., Ribodetti, A., Brossier, R., Operto, S. & Virieux, J. (2010a). Imaging of VTI media from wide-aperture data by frequency-domain full-waveform inversion. In *Expanded Abstracts, 72th Annual International Meeting, EAGE*.
- Gholami, Y., Ribodetti, A., Brossier, R., Operto, S. & Virieux, J. (2010b). Sensitivity analysis of full waveform inversion in VTI media. In *Expanded Abstracts, 72th Annual International Meeting, EAGE*, page 8192.
- Glagneaud, J. L. M. F. & Coppens, F. (1997). *Traitement du signal pour géologues et géophysiciens*. éditions TECHNIP.
- Gosselet, A. & Le Bégat, S. (2009). Combining borehole and surface seismic data for velocity field estimation through slope tomography. *Geophysical Journal International*, 176:897–908.

- Gouveia, W. P. & Scales, J. A. (1997). Resolution of seismic waveform inversion : Bayes versus occam. *Inverse Problems*, 13:323–349.
- Gouveia, W. P. & Scales, J. A. (1998). Bayesian seismic waveform inversion : parameter estimation and uncertainty analysis. *Journal of Geophysical Research*, 103(B2):2579–2779.
- Gray, S. H., Etgen, J., Dellinger, J. & Whitmore, D. (2001). Seismic migration problems and solutions. *Geophysics*, 66(5):1622–1640.
- Gregory, A. (1976). Fluid saturation effects on dynamic elastic properties of sedimentary rocks. *Geophysics*, 41:895–921.
- Guittou, A., Ayeni, G. & Diaz, E. (2012). Constrained full-waveform inversion by model reparameterization. *Geophysics*, 77(2):R117–R127.
- Guittou, A., Ayeni, G. & Gonzales, G. (2010). A preconditioning scheme for full waveform inversion. *SEG Technical Program Expanded Abstracts*, 29(1):1008–1012.
- Gutenberg, B. (1914). Über erdbebenwellen viia. beobachtungen an registrierungen von fernbeben in göttingen und folgerungen über die konstitution des erdkörpers. *Nachrichten von der Königlich Gesellschaft der Wissenschaften zu Göttinge, Mathematisch-Physikalische Klasse*, pages 125–176.
- Hadamard, J. (1902). Sur les problèmes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52.
- Hak, B. & Mulder, W. (2010). Migration for velocity and attenuation perturbations. *Geophysical Prospecting*, 58:939–951.
- Hak, B. & Mulder, W. A. (2011). Seismic attenuation imaging with causality. *Geophysical Journal International*, 184(1):439–451.
- Hale, D. (2011). Structure-oriented bilateral filtering of seismic images. *SEG Technical Program Expanded Abstracts*, 30(1):3596–3600.
- Hall, S. & Kendall, J. (2003). Fracture characterization at Valhall : Application of P-wave amplitude variation with offset and azimuth (AVOA) analysis to a 3D ocean-bottom data set. *Geophysics*, pages 1150–1160.
- Hall, S., Kendall, J. & Barkved, O. (2002). Fractured reservoir characterization using P-wave AVOA analysis of 3D OBC data. *The Leading Edge*, pages 777–781.
- Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4):561–580.
- Hatchell, P., Kwar, R. & Savitski, A. (2005). Integrating 4d seismic, geomechanics and reservoir simulation in the valhall oil field. *In Expanded Abstracts*.
- Heincke, B., Günther, T., Dalsegg, E., Ronning, J., Ganerod, G. & Elvebakk, H. (2009). Combined three-dimensional electric and seismic tomography study on the aknes rockslide in western norway. *Journal of Applied Geophysics*, page doi :10.1016/j.jappgeo.2009.12.004.

- Hicks, G. J. & Pratt, R. G. (2001). Reflection waveform inversion using local descent methods : estimating attenuation and velocity over a gas-sand deposit. *Geophysics*, 66(2):598–612.
- Hobro, J., Singh, S. & Minshull, T. (2002). Three-dimensional tomographic inversion of combined reflection and refraction seismic traveltimes data. *Geophysical Journal International*, 152(1):79–93.
- Hobro, J. W. D., Singh, S. C. & Minshull, T. A. (2003). Three-dimensional tomographic inversion of combined reflection and refraction seismic traveltimes data. *Geophysical Journal International*, 152:79–93.
- Hoven, J., Richardson, M. & Stoll, R., éditeurs (1991). *Assessment of shear strength of the sea bottom from shear wave velocity measurements on box cores and in-situ : Shear waves in marine sediments*. Shear waves in marine sediments : Kluwer Academic Publishers.
- Hu, W., Abubakar, A. & Habashy, T. (2009). Preconditioned non-linear conjugate gradient method for seismic full-waveform inversion. In *Expanded Abstracts*, page U018. EAGE.
- Huang, Y., MacBeth, C., Barkved, O. & van Gestel, J. (2011). Enhancing dynamic interpretation at the Valhall field by correlating well activity to 4D seismic signatures. *First Break*, 29.
- Hustedt, B., Operto, S. & Virieux, J. (2004). Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling. *Geophysical Journal International*, 157:1269–1296.
- Igel, H., Djikpesse, H. & Tarantola, A. (1996). Waveform inversion of marine reflection seismograms for p impedance and poisson’s ratio. *Geophysical Journal International*, 124:363–371.
- Igel, H. & Schoenberg, M. (1995). Shear property estimation using marine pressure and additional ocean bottom recordings. In *Extended Abstracts, 57th Annual EAGE Conference & Exhibition, Glasgow, UK*, page E040.
- Jacobs, G. & Hesthaven, J. S. (2006). High-order nodal discontinuous Galerkin particle-in-cell methods on unstructured grids. *Journal of Computational Physics*, 214:96–121.
- Jaiswal, P., Zelt, C., Dasgupta, R. & Nath, K. (2009). Seismic imaging of the Naga Thrust using multiscale waveform inversion. *Geophysics*, 74(6):WCC129–WCC140.
- Jamane, M., Beydoun, W., Crase, E., Cao, D., Koren, Z., Landa, E., Mendes, M., Pica, A., Noble, M., Roeth, G., Singh, S., Snieder, R., Tarantola, A. & Trezeguet, D. (1989). Wavelengths of Earth structures that can be resolved from seismic reflection data. *Geophysics*, 54(7):906–910.
- Jeong, W., Lee, H. & Min, D. (2012). Full waveform inversion strategy for density in the frequency domain. *Geophysical Journal International*, 188:1221–1242.
- Jin, S., Cambois, G. & Vuillermoz, C. (2000). Shear-wave velocity and density estimation from PS-wave AVO analysis : Application to an OBS dataset from the North Sea. *Geophysics*, 65:1446–1454.

- Jin, S., Madariaga, R., Virieux, J. & Lambaré, G. (1991). Two dimensional asymptotic iterative elastic inversion. *In Expanded Abstracts*, numéro SI5.6, pages 1013–1016. S.E.G.
- Jin, S., Madariaga, R., Virieux, J. & Lambaré, G. (1992). Two-dimensional asymptotic iterative elastic inversion. *Geophysical Journal International*, 108:575–588.
- Jo, C. H., Shin, C. & Suh, J. H. (1996). An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator. *Geophysics*, 61:529–537.
- Jones, T. (1986). Pore fluids and frequency-dependent wave propagation in rocks. *Geophysics*, 51:1939–1953.
- Kamei, R. & Pratt, R. G. (2008). Waveform tomography strategies for imaging attenuation structure for cross-hole data. *In Extended Abstracts, Roma*, page F019.
- Karypis, G. & Kumar, V. (1999). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1):359 – 392.
- Käser, M. & Dumbser, M. (2008). A highly accurate discontinuous Galerkin method for complex interfaces between solids and moving fluids. *Geophysics*, 73(3):23–35.
- Keho, T. H. (1988). Paraxial ray Kirchhoff migration. *Geophysics*, 53:1540–1546.
- Klimentos, T. & McCann, C. (1990). Relationships among compressional wave attenuation, porosity, clay content, and permeability in sandstones. *Geophysics*, 55(8):998–1014.
- Kolb, P., Collino, F. & Lailly, P. (1986). Prestack inversion of 1-D medium,. *In Extended Abstracts*, volume 74, pages 498–508.
- Kolsky, H. (1956). The propagation of stress pulses in viscoelastic solids. *Philosophical Magazine*, 1:693–710.
- Komatitsch, D., Barnes, C. & Tromp, J. (2000). Wave propagation near a fluid-solid interface : a spectral element approach. *Geophysics*, 65(2):623–631.
- Komatitsch, D. & Vilotte, J. P. (1998). The spectral element method : an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the Seismological Society of America*, 88:368–392.
- Kommedal, J. (2008). Use of descriptive statistics with repeatability data. SEG.
- Kommedal, J., abd J.P. Van Gestel, O. & Pettersen, R. (2007a). Processing strategies for multiple repeat 4d seismic. SEG.
- Kommedal, J. & Barkved, O. (2005). 4d repeatability beyond twice – using variance. pages 2526–2530. SEG.
- Kommedal, J., Semb, P. & Manning, T. (2007b). A case of si attenuation in 4d seismic data recorded with a permanently installed array. *Geophysics*, 72.
- Kommedal, J. H., Barkved, O. I. & Howe, D. J. (2004). Initial experience operating a permanent 4C seabed array for reservoir monitoring at Valhall. *SEG Technical Program Expanded Abstracts*, 23(1):2239–2242.

BIBLIOGRAPHIE

- Kommedal, J. H., Barkved, O. I. & Thomsen, L. A. (1997). Acquisition of 4 component obs data - a case study from the valhall field. *Presented at the 59th EAGE conference & Exhibition, Geneva, EAGE*.
- Kosloff, D., Kessler, D., Filho, A., Tessmer, E., Behle, A. & Strahilevitz, R. (1990). Solution of the equations of dynamic elasticity by a Chebychev spectral method. *Geophysics*, 55:464–473.
- Lailly, P. (1983a). The seismic inverse problem as a sequence of before stack migrations. In Bednar, R. & Weglein, éditeurs : *Conference on Inverse Scattering, Theory and application, Society for Industrial and Applied Mathematics, Philadelphia*, pages 206–220.
- Lailly, P. (1983b). The seismic problem as a sequence of before-stack migrations. In Bednar, J., éditeur : *Conference on Inverse Scattering : Theory and Applications*. SIAM, Philadelphia.
- Lambaré, G. (2008). Stereotomography. *Geophysics*, 73(5):VE25–VE34.
- Lambaré, G. & Alérini, M. (2005). Semi-automatic picking PP-PS stereotomography : application to the syntheticValhall dataset. In *Extended Abstracts*, pages 943–946.
- Lambaré, G., Alerini, M., Baina, R. & Podvin, P. (2004a). Stereotomography : a semi-automatic approach for velocity macromodel estimation. *Geophysical Prospecting*, 52:671–681.
- Lambaré, G., Alerini, M. & Podvin, P. (2004b). Stereotomographic picking in practice. In *Extended Abstracts, 74th annual meeting , (10-15 October 2004, Denver)*, page TOM 1.6. Soc. Expl. Geophys.
- Lambaré, G., Virieux, J., Madariaga, R. & Jin, S. (1992). Iterative asymptotic inversion in the acoustic approximation. *Geophysics*, 57:1138–1154.
- Lavergne, M. (1986). *Méthodes sismiques*. Editions Technip, 27 rue Ginoux, 75737 Paris Cédex 15.
- Le Bégat, S., Podvin, P. & Lambaré, G. (2000). Application of 2-D stereotomography to marine seismic reflection data. In *Expanded Abstracts, 70th Annual SEG Meeting and Exposition (6-11 August 2000, Calgary)*, pages 2142–2145. Soc. Expl. Geophys.
- Lee, H.-Y., Koo, J. M., Min, D.-J., Kwon, B.-D. & Yoo, H. S. (2010). Frequency-domain elastic full-waveform inversion for VTI media. *Geophysical Journal International*, 183(2):884–904.
- Leeuwen, T. V. & Mulder, W. A. (2010). A correlation-based misfit criterion for wave-equation travelttime tomography. *Geophysical Journal International*, 182(3):1383–1394.
- Lehmann, I. (1936). P'. *Publications du Bureau Central Séismologique International*, A14:87–115.
- Leonard, R. & Munns, J. (1987). *Valhall Field in Geology of Norwegian Oil and Gas Fields*. Graham and Trotman.
- Levander, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.
- LeVeque, R. J. (2002). *Finite volume methods for hyperbolic problems*. Cambridge University Press, New-York, USA.

- Liao, O. & McMechan, G. A. (1995). 2.5D full-wavefield viscoacoustic inversion. *Geophysical Prospecting*, 43:1043–1059.
- Liao, Q. & McMechan, G. A. (1996). Multifrequency viscoacoustic modeling and inversion. *Geophysics*, 61:1371–1378.
- Liu, Z. (1997). An analytical approach to migration velocity analysis. *Geophysics*, 62:1238–1249.
- Luo, Y. & Schuster, G. T. (1990). Parsimonious staggered grid finite-differencing of the wave equation. *Geophysical Research Letters*, 17(2):155–158.
- Luo, Y. & Schuster, G. T. (1991). Wave-equation travel time tomography. *Geophysics*, 56:645–653.
- Malinowski, M., Operto, S. & Ribodetti, A. (2011). High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion. *Geophysical Journal International*, 186(3):1179–1204.
- Manukyan, E., Latzel, S., Maurer, H., Marelli, S. & Greenhalgh, S. (2012). Exploitation of data-information content in elastic-waveform inversions. *Geophysics*, 77:105–115.
- Marfurt, K. (1984). Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49:533–549.
- Mari, J. L., Glangaud, F. & Coppens, F. (1999). *Signal processing for geologists & geophysicists*. éditions TECHNIP.
- Marquering, H., Dahlen, F. A. & Nolet, G. (1999). Three-dimensional sensitivity kernels for finite-frequency traveltimes : the banana-doughnut paradox. *Geophysical Journal International*, 137:805–815.
- Mavko, G., Mukerji, T. & Dvorkin, J. (2009). *The Rocks Physics Handbooks, Tools for Seismic Analysis in Porous Media ; Second Edition*. Cambridge University Press, Cambridge, UK.
- McMechan, G. (1989). A review of seismic acoustic imaging by reverse time migration. *International Journal of Imaging Systems and Technology*, 1:18–21.
- Menke, W. (1984). *Geophysical Data Analysis : Discrete Inverse Theory*. Academic Press, Inc., Orlando, USA.
- Métivier, L., Brossier, R., Virieux, J. & Operto, S. (2012a). Toward gauss-newton and exact newton optimization for full waveform inversion. In *EAGE, 74th Conference and Exhibition*, page This issue.
- Métivier, L., Halpern, L., Delprat-Jannaud, F. & Lailly, P. (2012b). A 2D non linear inversion of well seismic data. *Inverse Problems*, 27(5):055005 doi :10.1088/0266-5611/27/5/055005.
- Meyerholtz, K. A., Pavlis, G. L. & Szpakowski, S. A. (1989). Convolutional quelling in seismic tomography. *Geophysics*, 54(5):570–580.

- Meza-Fajardo, K. & Papageorgiou, A. (2008). A nonconvolutional, split-field, perfectly matched layer for wave propagation in isotropic and anisotropic elastic media : Stability analysis. *Bulletin of the Seismological Society of America*, 98(4):1811–1836.
- Miller, D., Oristaglio, M. & Beylkin, G. (1987). A new slant on seismic imaging : Migration and integral geometry. *Geophysics*, 52(7):943–964.
- Min, D. J. & Shin, C. (2006). Refraction tomography using a waveform-inversion back-propagation technique. *Geophysics*, 71(3):R21–R30.
- Min, D.-J., Shin, C., Pratt, R. G. & Yoo, H. S. (2003). Weighted-averaging finite-element method for 2D elastic wave equations in the frequency domain. *Bulletin of the Seismological Society of America*, 93(2):904–921.
- Moczo, P., Kristek, J. & Halada, L. (2004). *The Finite-Difference Method for Seismologists An Introduction*. Comenius University Bratislava.
- Mora, P. R. (1987). Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, 52:1211–1228.
- Mora, P. R. (1988). Elastic wavefield inversion of reflection and transmission data. *Geophysics*, 53:750–759.
- Mordret, A., Shapiro, N., Singh, S., P. Roux & Barkved, O. (2011). Low-frequency passive seismic tomography using valhall lofs. *In Expanded Abstracts*.
- Morency, C., Luo, Y. & Tromp, J. (2011). Acoustic, elastic and poroelastic simulations of CO_2 sequestration crosswell monitoring based on spectral-element and adjoint methods. *Geophysical Journal International*, 185:955–966.
- Moritz, H. & Sunkel, H. (1978). *Approximation methods in geodesy*. H. Wichmann, Karlsruhe.
- Mueller, M., Barkved, O. & Thomsen, L. (1997). Dipole sonic results (valhall area) - implications for avo and obs interpretation. *In Extended Abstracts, 59th Annual EAGE Conference & Exhibition, Geneva, Switzerland*.
- Mulder, W. & Hak, B. (2009). An ambiguity in attenuation scattering imaging. *Geophysical Journal International*, 178:1614–1624.
- MUMPS-team (2006). *MUMPS - Multifrontal Massively Parallel Solver users' guide - version 4.6.3 (February 2006)*. ENSEEIHT-ENS Lyon, <http://www.enseeiht.fr/apo/MUMPS/> or <http://graal.ens-lyon.fr/MUMPS>.
- MUMPS-team (2009). *MUMPS - Multifrontal Massively Parallel Solver users' guide - version 4.9.2 (November 5, 2009)*. ENSEEIHT-ENS Lyon, <http://www.enseeiht.fr/apo/MUMPS/> or <http://graal.ens-lyon.fr/MUMPS>.
- Munns, J. W. (1985). The Valhall field : a geological overview. *Marine and Petroleum Geology*, 2:23–43.
- Muyzert, E. (2000). Scholte wave velocity inversion for a near surface s-velocity model and ps-statics. *In Expanded Abstracts, 70 Annual SEG Meeting and Exposition*. Society of Exploration Geophysics.

- Nag, S., Alerini, M., Duvencak, E. & Ursin, B. (2006). 2-D stereotomography for anisotropic media. *In Expanded Abstracts*, pages 3305–3308.
- Nag, S., Alerini, M. & Ursin, B. (2010). PP/PS anisotropic stereotomography. *Geophysical Journal International*, 181(1):427–452.
- Neves, F. A. & Singh, S. C. (1996). Sensitivity study of seismic reflection/refraction data. *Geophysical Journal International*, 126(2):470–476.
- Nihei, K. T. & Li, X. (2007). Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD). *Geophysical Journal International*, 169:1069–1078.
- Nocedal, J. (1980). Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation*, 35(151):773–782.
- Nocedal, J. & Wright, S. J. (1999). *Numerical Optimization*. New York, US : Springer.
- O. Barkved, P. H. (2003). Society of petroleum engineers inc. valhall field - still on plateau after 20 years of production. *This paper was prepared for presentation at Offshore Europe 2003 held in Aberdeen, UK*.
- O'Brien, M., Whitmore, N., Brandsberg-Dahl, S., Edgen, J. & Murphy, G. (1999). Multicomponent modelling of the Valhall field. *In Expanded Abstracts*. EAGE.
- Oldham, R. (1906). The constitution of the earth. *Quarterly Journal of the Geological Society of London*, 62:456–475.
- Olofsson, B., Probert, T., Kommedal, J. & Barkved, O. (2003). Azimuthal anisotropy from the valhall 4C 3D survey. *The Leading Edge*, pages 1228–1235.
- Operto, S. (2006). Frequency-domain full-waveform inversion of global offset data. Rapport technique, UMR Géosciences Azur - CNRS. Ecole thématique SEISCOPE (<http://seiscope.oca.eu/spip.php?rubrique20>).
- Operto, S. & Ph.Charvis (1996). Deep structure of the southern Kerguelen plateau (southern Indian ocean) from ocean bottom seismometer wide-angle seismic data. *Journal of Geophysical Research*, 101:25077–25103.
- Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A. & Dell'Aversana, P. (2004). Quantitative imaging of complex structures from multi-fold wide aperture seismic data. *Geophysical Prospecting*, 52:625–651.
- Operto, S., Virieux, J., Dessa, J. X. & Pascal, G. (2006). Crustal imaging from multi-fold ocean bottom seismometers data by frequency-domain full-waveform tomography : application to the eastern Nankai trough. *Journal of Geophysical Research*, 111(B09306): doi :10.1029/2005JB003835.
- Operto, S., Virieux, J., Ribodetti, A. & Anderson, J. E. (2009). Finite-difference frequency-domain modeling of visco-acoustic wave propagation in two-dimensional TTI media. *Geophysics*, 74 (5):T75–T95.

- Operto, S., Xu, S. & Lambaré, G. (2000). Can we image quantitatively complex models with rays? *Geophysics*, 65(4):1223–1238.
- Paige, C. C. & Saunders, M. A. (1982). ALGORITHM 583 LSQR : Sparse linear equations and least squares problems. *ACM Transactions on Mathematical Software*, 8(2):195–209.
- Pica, A., Diet, J. P. & Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in laterally invariant medium. *Geophysics*, 55(3):284–292.
- Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Plessix, R. E. (2009a). Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, 74(6):WCC53–WCC61.
- Plessix, R. E. (2009b). Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, 74:WCC149–WCC157.
- Plessix, R. E., Baeten, G., de Maag, J. W., Klaassen, M., Rujie, Z. & Zhifei, T. (2010a). Application of acoustic full waveform inversion to a low-frequency large-offset land data set. *SEG Technical Program Expanded Abstracts*, 29(1):930–934.
- Plessix, R. E. & Cao, Q. (2011). A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*.
- Plessix, R. E., Chavent, G. & De Roeck, Y. (1995). Automatic and simultaneous migration velocity analysis and waveform inversion of real data using a MBTT/WKBJ formulation. *In Expanded Abstracts*, pages 1099–1101. Society of Exploration Geophysics.
- Plessix, R. E., Michelet, S., Rynja, H., Kuehl, H., Perkins, C., de Maag, J. W. & Hatchell, P. (2010b). Some 3d applications of full waveform inversion. *In Expanded Abstracts*. EAGE.
- Plessix, R. E. & Perkins, C. (2010). Full waveform inversion of a deep water ocean bottom seismometer dataset. *First Break*, 28:71–78.
- Podvin, P. & Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity model : a massively parallel approach and its associated tools. *Geophysical Journal International*, 105:271–284.
- Polak, E. & Ribière, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue Française d'Informatique et de Recherche Opérationnelle*, 16:35–43.
- Pratt, R. G. (1990). Frequency-domain elastic modeling by finite differences : a tool for crosshole seismic imaging. *Geophysics*, 55(5):626–632.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part I : theory and verification in a physic scale model. *Geophysics*, 64:888–901.
- Pratt, R. G. (2008). Waveform tomography - successes, cautionary tales, and future directions. *In Presented at the 70th Annual EAGE Conference & Exhibition, Roma*, pages WO11 – Full-Waveform Inversion : current status and perspectives.

- Pratt, R. G., Hou, F., Bauer, K. & Weber, M. (2005). *Waveform tomography images of velocity and inelastic attenuation from the Mallik 2002 crosshole seismic surveys*. Geological survey of Canada.
- Pratt, R. G., Plessix, R. E. & Mulder, W. A. (2001). Seismic waveform tomography : the effect of layering and anisotropy. *In Presented at the 63th EAGE Conference & Exhibition, Amsterdam, EAGE.*, page P092.
- Pratt, R. G. & Sams, M. S. (1996). Reconciliation of crosshole seismic velocities with well information in a layered sedimentary environment. *Geophysics*, 61:549–560.
- Pratt, R. G., Shin, C. & Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.
- Pratt, R. G. & Shipp, R. M. (1999). Seismic waveform inversion in the frequency domain, part II : Fault delineation in sediments using crosshole data. *Geophysics*, 64:902–914.
- Pratt, R. G., Song, Z. M. & Warner, M. (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, 124:323–340.
- Pratt, R. G. & Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I : acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.
- Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. & Kommedal, J. (2011). On the footprint of anisotropy on isotropic full waveform inversion : the Valhall case study. *Geophysical Journal International*, 187:1495–1515.
- Prieux, V., Lambar/e, G., Operto, S. & Virieux, J. (2012). Building starting model for full waveform inversion from wide-aperture data by stereotomography. *Geophysical Prospecting*, in-press.
- Prieux, V., Operto, S., Brossier, R. & Virieux, J. (2009). Application of acoustic full waveform inversion to the synthetic Valhall model. volume 28, pages 2268–2272. SEG.
- Prieux, V., Operto, S., Lambaré, G. & Virieux, J. (2010). Building starting model for full waveform inversion from wide-aperture data by stereotomography. *SEG Technical Program Expanded Abstracts*, 29(1):988–992.
- Rao, Y. & Wang, Y. (2008). The strategies for attenuation inversion with waveform tomography. *In 70th Annual Conference. EAGE.*
- Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A. & dell’Aversana, P. (2004). Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions : application to a thrust belt. *Geophysical Journal International*, 159:1032–1056.
- Ribodetti, A., Operto, S., Virieux, J., Lambaré, G., Valéro, H.-P. & Gibert, D. (2000a). Asymptotic viscoacoustic diffraction tomography of ultrasonic laboratory data : a tool for rock properties analysis. *Geophysical Journal International*, 140:324–340.

- Ribodetti, A., Thierry, P., Lambaré, G. & Operto, S. (2000b). Improved multiparameter ray+born migration/inversion. In *Expanded Abstracts*, pages 1032–1035. Society of Exploration Geophysics.
- Ribodetti, A. & Virieux, J. (1996). Asymptotic theory for imaging the attenuation factors Q_p and Q_s . In *Inverse Problems of Wave Propagation and Diffraction, Proceedings, Aix-les-Bains, France 1996*, pages 334–353. Springer-Verlag.
- Ribodetti, A. & Virieux, J. (1998). Asymptotic theory for imaging the attenuation factor Q . *Geophysics*, 64:1–12.
- Ridder, S. D. & Dellinger, J. (2011). Ambient seismic noise eikonal tomography for near-surface imaging at valhall. *The Leading Edge*, 30:506–512.
- Robertson, J. O. A., Holliger, K., Green, A. G., Pugin, A. & Iaco, R. D. (1996). Effects of near-surface waveguides on shallow high-resolution seismic refraction and reflection data. *Geophysical Research Letters*, 23(5):495–498.
- Rosland, B., Tree, E., Barkved, O. & Kristiansen, P. (1999). Acquisition of 3d/4c obs data at valhall. In *61th Annual Conference*. EAGE.
- Saad, Y. (2003). *Iterative methods for sparse linear systems*. SIAM, Philadelphia.
- Sambridge, M. & Drijkoningen, G. (1992). Genetic algorithms in seismic waveform inversion. *Geophysical Journal International*, 109:323–342.
- Sambridge, M. S., Tarantola, A. & Kennett, B. L. (1991). An alternative strategy for non-linear inversion of seismic waveforms. *Geophysical Prospecting*, 39:723–736.
- Sams, M., Neep, J. & Worthington, M. (1997). The measurement of velocity dispersion and frequency-dependent intrinsic attenuation in sedimentary rocks. *Geophysics*, 62:1456–1464.
- Sayers, C. (1994). The elastic anisotropy of shales. *Journal of Geophysical Research*, 99:767–774.
- Sears, T., Singh, S. & Barton, P. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting*, 56(6):843–862.
- Sears, T. J., Barton, P. J. & Singh, S. C. (2010). Elastic full waveform inversion of multicomponent ocean-bottom cable seismic data : Application to alba field, u. k. north sea. *Geophysics*, 75(6):R109–R119.
- Shen, P. & Symes, W. W. (2008). Automatic velocity analysis via shot profile migration. *Geophysics*, 73(5):VE49–VE59.
- Sheng, J., Leeds, A., Buddensiek, M. & Schuster, G. T. (2006). Early arrival waveform tomography on near-surface refraction data. *Geophysics*, 71(4):U47–U57.
- Sheriff, R. (1975). Factors affecting seismic amplitudes. *Geophysical Prospecting*, 23.
- Sheriff, R. E. & Geldart, L. P. (1995). *Exploration seismology, second edition*. Cambridge University Press.

- Shi, Y., Zhao, W. & Cao, H. (2007). Nonlinear process control of wave-equation inversion and its application in the detection of gas. *Geophysics*, 72(1):R9–R18.
- Shin, C. & Cha, Y. H. (2008). Waveform inversion in the Laplace domain. *Geophysical Journal International*, 173(3):922–931.
- Shin, C. & Cha, Y. H. (2009). Waveform inversion in the Laplace-Fourier domain. *Geophysical Journal International*, 177:1067–1079.
- Shin, C. & Ha, W. (2008). A comparison between the behavior of objective functions for waveform inversion in the frequency and laplace domains. *Geophysics*, 73(5):VE119–VE133.
- Shin, C., Jang, S. & Min, D. J. (2001). Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, 49:592–606.
- Shiobara, H., Sugioka, H., Mochizuki, K., Oki, S., Kanazawa, T., Fukao, Y. & Suyehiro, Y. (2010). Double seismic zone in the north mariana region revealed by long-term ocean bottom array observation. *Geophysical Journal International*, 183:1455–1469.
- Shipp, R. M. & Singh, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151:325–344.
- Sirgue, L. (2003). *Inversion de la forme d'onde dans le domaine fréquentiel de données sismiques grand offset*. Thèse de doctorat, Université Paris 11, France - Queen's University, Canada.
- Sirgue, L. (2006). The importance of low frequency and large offset in waveform inversion. *In Presented at the 68th EAGE Conference & Exhibition, Vienna, EAGE*, page A037.
- Sirgue, L., Barkved, O. I., Dellinger, J., Etgen, J., Albertin, U. & Kommedal, J. H. (2010). Full waveform inversion : the next leap forward in imaging at Valhall. *First Break*, 28:65–70.
- Sirgue, L., Barkved, O. I., Gestel, J. P. V., Askim, O. J. & Kommedal, J. H. . (2009). 3D waveform inversion on Valhall wide-azimuth OBC. *In Presented at the 71th Annual International Meeting, EAGE, Expanded Abstracts*, page U038.
- Sirgue, L., Denel, B. & Gao, F. (2011). Integrating 3d full waveform inversion into depth imaging projects. *SEG Technical Program Expanded Abstracts*, 30(1):2354–2358.
- Sirgue, L., Etgen, J. T. & Albertin, U. (2008). 3D Frequency Domain Waveform Inversion using Time Domain Finite Difference Methods. *In Proceedings 70th EAGE, Conference and Exhibition, Roma, Italy*, page F022.
- Sirgue, L. & Pratt, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.
- Smithyman, B., Pratt, R. G., Hayles, J. & Wittebolle, R. (2009). Detecting near-surface objects with seismic waveform tomography. *Geophysics*, 74(6):WCC119–WCC127.
- Song, Z. & Williamson, P. (1995). Frequency-domain acoustic-wave modeling and inversion of crosshole data. part 1 : 2.5-D modeling method. *Geophysics*, 60(3):784–795.

- Song, Z., Williamson, P. & Pratt, G. (1995). Frequency-domain acoustic-wave modeling and inversion of crosshole data. Part 2 : Inversion method, synthetic experiments and real-data results. *Geophysics*, 60(3):786–809.
- Soubaras, R. & Whiting, P. (2011). Variable depth streamer — the new broadband acquisition system. *SEG Technical Program Expanded Abstracts*, 30(1):4349–4353.
- Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y. (2009a). FWT2D : A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—Part 1 : Algorithm. *Computers & Geosciences*, 35(3):487 – 495.
- Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y. (2009b). FWT2D : A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—Part 2 : Numerical examples and scalability analysis. *Computers & Geosciences*, 35(3):496 – 514.
- Stewart, P. (2006). Interferometric imaging of ocean bottom noise. volume 25, pages 1555–1559. SEG.
- Stopin, A. & Ehinger, A. (2001). Joint PP PS tomographic inversion of the Mahogany 2D-4C OBC seismic data. *In Expanded Abstracts*, pages 837–840. Society of Exploration Geophysics.
- Sun, R. & McMechan, G. (1992). 2-d full wavefield inversion for wide-aperture, elastic, seismic data. *Geophysical Journal International*, 111:1–10.
- Symes, W. W. (1993). A differential semblance criterion for inversion of multioffset seismic reflection data. *Journal of Geophysical Research*, 98:2061–2073.
- Symes, W. W. (2008). Migration velocity analysis and waveform inversion. *Geophysical Prospecting*, 56:765–790.
- Symes, W. W. & Carazzone, J. J. (1991). Velocity inversion by differential semblance optimization. *Geophysics*, 56:654–663.
- Taillandier, C., Noble, M., Chauris, H. & Calandra, H. (2009). First-arrival travel time tomography based on the adjoint state method. *Geophysics*, 74(6):WCB1–WCB10.
- Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tarantola, A. (1984b). Linearized inversion of seismic reflection data. *Geophysical Prospecting*, 32:998–1015.
- Tarantola, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.
- Tarantola, A. (1987). *Inverse problem theory : methods for data fitting and model parameter estimation*. Elsevier, New York.
- Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms including elasticity and attenuation. *Pure and Applied Geophysics*, 128:365–399.

- Tarantola, A. (2005). *Inverse Problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia.
- Tatham, R. & Stoffa, R. (1976). Vp/vs - a potential hydrocarbon indicator. *Geophysics*, 41:837–849.
- Thierry, P. & Baina, R. (2001). 3-D PAPsDM based on ray+Kirchhoff approximation : Theory and callibration. *In DIG annual report*, volume 1, pages 65–90. Depth Imaging Group.
- Thierry, P., Lambaré, G., Podvin, P. & Noble, M. (1999a). 3-D preserved amplitude prestack depth migration on a workstation. *Geophysics*, 64(1):222–229.
- Thierry, P., Operto, S. & Lambaré, G. (1999b). Fast 2D ray-Born inversion/migration in complex media. *Geophysics*, 64(1):162–181.
- Thomsen, L. A. (1986). Weak elastic anisotropy. *Geophysics*, 51:1954–1966.
- Thomsen, L. A., Barkved, O., Haggard, B., Kommedal, J. H. & Rosland, B. (1997). Converted wave imaging of valhall reservoir. *Presented at the 59th EAGE conference & Exhibition, Geneva, EAGE*.
- Toksöz, M. N. & Johnston, D. H. (1981). *Geophysics reprint series, No. 2 : Seismic wave attenuation*. Society of exploration geophysicists, Tulsa, OK.
- Tonellot, T. & Broadhead, M. (2011). Stratigraphic inversion with anisotropic diffusion regularization. *In EAGE, 73rd Conference and Exhibition*, page P312.
- Tonn, R. (1991). The determination of seismic quality factor q from vsp data : A comparison of different computational methods. *Geophysical Prospecting*, 39:1–27.
- Toomey, D. R., Solomon, S. C. & Purdy, G. M. (1994). Tomographic imaging of the shallow crustal structure of the east pacific rise at 9deg.30'N. *Journal of Geophysical Research*, 99:24135–24157.
- Toverud, T. & Ursin, B. (2005). Comparison of seismic attenuation models using zero-offset vertical seismic profiling (vsp) data. *Geophysics*, 70:F17–F25.
- Tsvankin, I. (1995). Normal moveout from dipping reflectors in anisotropic media. *Geophysics*, 60:268–284.
- Tsvankin, I. (2001). Seismic signature and analysis of reflection data in anisotropic media. *In* Helbig, K. & Treitel, S., éditeurs : *Seismic Exploration*, volume 29 de *Handbook of Geophysical Exploration*. Pergamon.
- van den Berg, P., Abubakar, A. & Fokkema, J. (2003). Multiplicative regularization for contrast profile inversion. *Radio Science*, 38:23.1–23.10.
- van Gestel, J., Best, K., Barkved, O. & Kommedal, J. (2008a). Integrating frequent time-lapse data into the reservoir simulation modeling of the valhall field. *In Expanded Abstracts*.
- van Gestel, J., Best, K. D., Barkved, O. I. & Kommedal, J. H. (2011). Integration of the Life of Field Seismic data with the reservoir model at the Valhall field. *Geophysical Prospecting*, 59:673–681.

- van Gestel, J., Kommedal, J., Barkved, O., Mundal, I., Bakke, R. & Best, K. (2008b). Continuous seismic surveillance of Valhall field. *The Leading Edge*, pages 1616–1621.
- Vanorio, T. & Virieux, J. (2005). Three-dimensional seismic tomography from p wave and s wave microearthquake travel times and rock physics characterization of the campi flegrei caldera. *Journal of Geophysical Research*, 110.
- Vigh, D., Kapoor, J., Moldoveanu, N. & Li, H. (2011). Breakthrough acquisition and technologies for subsalt imaging. *Geophysics*, 76(5):WB41–WB51.
- Vigh, D., Starr, B., Kapoor, J. & Li, H. (2010). 3d full waveform inversion on a gulf of mexico waz data set. *SEG Technical Program Expanded Abstracts*, 29(1):957–961.
- Virieux, J. (1984). SH wave propagation in heterogeneous media, velocity-stress finite difference method. *Geophysics*, 49:1259–1266.
- Virieux, J. (1986a). P-SV wave propagation in heterogeneous media, velocity-stress finite difference method. *Geophysics*, 51:889–901.
- Virieux, J. (1986b). P-SV wave propagation in heterogeneous media : velocity stress finite difference method. *Geophysics*, 51:889–901.
- Virieux, J. & Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC127–WCC152.
- Wang, B. (1993). *Improvement of seismic travel-time inversion methods and application to observed data*. Thèse de doctorat, Purdue University.
- Wang, Y. (2008). *Seismic Inverse Q Filtering*. Wiley-Blackwell.
- Wang, Y. & Rao, Y. (2009). Reflection seismic waveform tomography. *Journal of Geophysical Research*, 114(B03304):doi :10.1029/2008JB005916.
- Wapenaar, C., Verschuur, E. & Herrmann, P. (1992). Amplitude preprocessing of single- and multi-component seismic data. *Geophysics*, 57:1178—1188.
- Waters, K. (1978). *Reflection Seismology*. John Wiley and Sons, Inc.
- White, R. S. & Stephen, R. A. (1980). Compressional to shear wave conversion in oceanic crust. *Geophysical Journal of the Royal Astronomical Society*, 63:547–565.
- Williamson, P. (1991). A guide to the limits of resolution imposed by scattering in ray tomography. *Geophysics*, 56:202–207.
- Williamson, P. & Pratt, G. (1995). A critical review of 2.5D acoustic wave modeling procedures. *Geophysics*, 60:591–595.
- Winkler, K. & Nur, A. (1982). Seismic attenuation : effects of pore fluids and frictional sliding. *Geophysics*, 47:1–15.
- Woodward, M. J. (1992). Wave-equation tomography. *Geophysics*, 57:15–26.
- Woodward, M. J., Nichols, D., Zdraveva, O., Whitfield, P. & Johns, T. (2008). A decade of tomography. *Geophysics*, 73(5):VE5–VE11.

- Wu, R. S. & Aki, K. (1985). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, 50(4):582–595.
- Wu, R. S. & Toksöz, M. N. (1987). Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, 52:11–25.
- Yang, T. & Sava, P. (2011a). Image-domain waveform tomography with two-way wave-equation. *SEG Technical Program Expanded Abstracts*, 30(1):2591–2596.
- Yang, T. & Sava, P. (2011b). Wave-equation migration velocity analysis with time-shift imaging. *Geophysical prospecting*, 59(4):635–650.
- Yilmaz, Ö. (1987). *Seismic data processing*. Society of Exploration Geophysicists.
- Yilmaz, Ö. & Chambers, R. (1984). Migration velocity analysis by wavefield extrapolation. *Geophysics*, 49:1664–1674.
- Zelt, C. & Barton, P. J. (1998). Three-dimensional seismic refraction tomography : a comparison of two methods applied to data from the Faeroe basin. *Journal of Geophysical Research*, 103(B4):7187–7210.
- Zelt, C. & Smith, R. B. (1992). Seismic travelttime inversion for 2-D crustal velocity structure. *Geophysical Journal International*, 108:16–34.
- Zhou, Y., Dahlen, F., Nolet, G. & Laske, G. (2005). Finite-frequency effects in global surface-wave tomography. *Geophysical Journal International*, 163:1087–1111.
- Zhu, C., Byrd, R. H. & Nocedal, J. (1997). L-bfgs-b : Algorithm 778 : L-bfgs-b, fortran routines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560.
- Zienkiewicz, O. C., Taylor, R. L. & Zhu, J. Z. (2005). *The Finite Element Method : Its Basis and Fundamentals*. Elsevier, London. 6th edition.

Annexe A

Implémentation de la paramétrisation (V_P, I_P, V_S)

Nous voulons montrer dans cette annexe d'où proviennent les relations utilisées pour l'implémentation des familles de paramètre dans notre algorithme d'inversion, en prenant l'exemple de la famille V_P, I_P, V_S). Pour inverser un des paramètres appartenant à cette famille, il faut fournir au code les modèles de vitesse V_P et V_S , ainsi que la densité de la même manière que si l'on considérait la famille (V_P, ρ, V_S) . Le modèle de densité fourni est cependant dérivé de la division du modèle initial d'impédance désiré par le modèle de vitesse initial. Notons que dans le cas où l'impédance est inversée en utilisant la régularisation Tikhonov, il faut aussi fournir un modèle a priori d'impédance et un modèle d'amortissement σ_{I_P} à l'ordre de grandeur approprié. Les modèles initiaux de vitesses et de densité sont ensuite transformés dans la famille (λ, μ, ρ) afin de calculer les champs d'onde incidents et rétropropagés. La formule du gradient pour le paramètre V_P dans la famille de paramètre (V_P, I_P, V_S) est donnée par :

$$\mathbf{G}_{V_P} = \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{u}^t \left(\frac{\partial \mathbf{A}^t}{\partial V_P} \right)_{I_P, V_S} \mathbf{A}^{-1} \mathbf{W}_d \Delta \mathbf{d}^* \right\}, \quad (\text{A.1})$$

où la dérivée partielle de \mathbf{A} par rapport à V_P se fait à I_P et V_S constants. Il nous faut l'expression de cette dérivée partielle en terme de quantités exprimées dans la paramétrisation (λ, μ, ρ) , qui est la formulation utilisée dans le problème direct. Ce changement de paramétrisation se fait via un développement en chaîne de la dérivée partielle en introduisant les dépendances existant entre ces deux paramétrisations :

$$\left(\frac{\partial \mathbf{A}^t}{\partial V_P} \right)_{I_P, V_S} = \left(\frac{\partial \mathbf{A}^t}{\partial \lambda} \right)_{\mu, \rho} \cdot \left(\frac{\partial \lambda}{\partial V_P} \right)_{I_P, V_S} + \left(\frac{\partial \mathbf{A}^t}{\partial \mu} \right)_{\lambda, \rho} \cdot \left(\frac{\partial \mu}{\partial V_P} \right)_{I_P, V_S} + \left(\frac{\partial \mathbf{A}^t}{\partial \rho} \right)_{\lambda, \mu} \cdot \left(\frac{\partial \rho}{\partial V_P} \right)_{I_P, V_S} \quad (\text{A.2})$$

Le développement de ce calcul nécessite de rappeler d'une part, les relations existant entre les paramètres de Lamé et les vitesses :

$$\begin{aligned} \lambda &= \rho V_P^2 - 2\rho V_S^2 = I_P V_P - 2 \frac{I_P}{V_P} V_S^2 \\ \mu &= \rho V_S^2 = \frac{I_P}{V_P} V_S^2 \\ \rho &= \frac{I_P}{V_P}, \end{aligned} \quad (\text{A.3})$$

et d'autres part l'expression du système elliptique d'ordre 2 en vitesse employé dans le problème direct isotrope :

$$\begin{aligned}
 \omega^2 \rho(\mathbf{x}) V_x(\mathbf{x}, \omega) &= -\frac{\partial}{\partial x} \left\{ (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_x(\mathbf{x}, \omega)}{\partial x} + \lambda(\mathbf{x}) \frac{\partial V_z(\mathbf{x}, \omega)}{\partial z} \right\} - \\
 &\quad \frac{\partial}{\partial z} \left\{ \mu(\mathbf{x}) \left(\frac{\partial V_x(\mathbf{x}, \omega)}{\partial z} + \frac{\partial V_z(\mathbf{x}, \omega)}{\partial x} \right) \right\} + f_x(\mathbf{x}, \omega) \\
 \omega^2 \rho(\mathbf{x}) V_z(\mathbf{x}, \omega) &= -\frac{\partial}{\partial x} \left\{ \mu(\mathbf{x}) \left(\frac{\partial V_x(\mathbf{x}, \omega)}{\partial z} + \frac{\partial V_z(\mathbf{x}, \omega)}{\partial x} \right) \right\} - \\
 &\quad \frac{\partial}{\partial z} \left\{ \lambda(\mathbf{x}) \frac{\partial V_x(\mathbf{x}, \omega)}{\partial x} + (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_z(\mathbf{x}, \omega)}{\partial z} \right\} + f_z(\mathbf{x}, \omega)
 \end{aligned} \tag{A.4}$$

Ce système d'équation peut s'écrire sous la forme : $\mathbf{A} \mathbf{p} = \mathbf{s}$, c'est à dire :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} V_x(\mathbf{x}, \omega) \\ V_z(\mathbf{x}, \omega) \end{pmatrix} = \begin{pmatrix} f_x(\mathbf{x}, \omega) \\ f_z(\mathbf{x}, \omega) \end{pmatrix} \tag{A.5}$$

où l'on a :

$$\begin{aligned}
 a_{11} &= \omega^2 \rho + \frac{\partial}{\partial x} \left\{ (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial(\cdot)}{\partial x} + \right\} + \frac{\partial}{\partial z} \left\{ \mu(\mathbf{x}) \frac{\partial(\cdot)}{\partial z} \right\} \\
 a_{12} &= \frac{\partial}{\partial x} \left\{ \lambda \frac{\partial(\cdot)}{\partial z} \right\} + \frac{\partial}{\partial z} \left\{ \mu \frac{\partial(\cdot)}{\partial x} \right\} \\
 a_{21} &= \frac{\partial}{\partial x} \left\{ \mu \frac{\partial(\cdot)}{\partial z} \right\} + \frac{\partial}{\partial z} \left\{ \lambda \frac{\partial(\cdot)}{\partial x} \right\} \\
 a_{22} &= \omega^2 \rho + \frac{\partial}{\partial z} \left\{ (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial(\cdot)}{\partial z} + \right\} + \frac{\partial}{\partial x} \left\{ \mu(\mathbf{x}) \frac{\partial(\cdot)}{\partial x} \right\}
 \end{aligned} \tag{A.6}$$

Les relations (A.3) nous donnent :

$$\begin{aligned}
 \left(\frac{\partial \lambda}{\partial V_P} \right)_{I_P, V_S} &= I_P + 2 \frac{I_P V_S^2}{V_P^2} \\
 \left(\frac{\partial \mu}{\partial V_P} \right)_{I_P, V_S} &= -\frac{I_P V_S^2}{V_P^2} \\
 \left(\frac{\partial \rho}{\partial V_P} \right)_{I_P, V_S} &= -\frac{I_P}{V_P^2}
 \end{aligned} \tag{A.7}$$

et le système (A.4) permet d'obtenir :

$$\begin{aligned}
 \left(\frac{\partial \mathbf{A}^t}{\partial \lambda} \right)_{\mu, \rho} &= \mathfrak{D}_\lambda = \begin{pmatrix} \partial_x \partial_x & \partial_x \partial_z \\ \partial_z \partial_x & \partial_z \partial_z \end{pmatrix} \\
 \left(\frac{\partial \mathbf{A}^t}{\partial \mu} \right)_{\lambda, \rho} &= \mathfrak{D}_\mu = \begin{pmatrix} \partial_x 2\partial_x + \partial_z \partial_z & \partial_z \partial_x \\ \partial_x \partial_z & \partial_z 2\partial_z + \partial_x \partial_x \end{pmatrix} \\
 \left(\frac{\partial \mathbf{A}^t}{\partial \rho} \right)_{\lambda, \mu} &= \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega^2 \end{pmatrix}
 \end{aligned} \tag{A.8}$$

où \mathfrak{D} représente l'opérateur des dérivées spatiales relatif au schéma de discrétisation employé. On peut ainsi réécrire la matrice du diagramme de rayonnement comme

$$\left(\frac{\partial \mathbf{A}^t}{\partial V_P}\right)_{I_P, V_S} = \mathfrak{D}_\lambda \left(I_P + 2\frac{I_P V_S^2}{V_P^2}\right) - \mathfrak{D}_\mu \left(\frac{I_P V_S^2}{V_P^2}\right) - \omega^2 \frac{I_P}{V_P^2} \mathbf{I} \quad (\text{A.9})$$

Le terme $-\omega^2 \frac{I_P}{V_P^2}$ est appelé terme de collocation en raison de sa localisation sur la diagonale de la matrice de rayonnement. On peut obtenir de la même manière les matrices de rayonnement dans cette paramétrisation pour l'impédance et V_s :

$$\begin{aligned} \left(\frac{\partial \mathbf{A}^t}{\partial I_P}\right)_{V_P, V_S} &= \mathfrak{D}_\lambda \left(V_P - 2\frac{V_S^2}{V_P}\right) + \mathfrak{D}_\mu \left(\frac{V_S^2}{V_P}\right) + \omega^2 \frac{1}{V_P} \mathbf{I} \\ \left(\frac{\partial \mathbf{A}^t}{\partial V_S}\right)_{V_P, I_P} &= \mathfrak{D}_\lambda \left(-4\frac{I_P V_S}{V_P}\right) + \mathfrak{D}_\mu \left(2\frac{I_P V_S}{V_P}\right) \end{aligned} \quad (\text{A.10})$$

Une fois les matrices de rayonnement calculées pour les paramètres inversés dans la paramétrisation choisie, on dérive leurs gradients et leurs modèles de perturbation. À l'issue de cette itération, les paramètres inversés sont à nouveau transformés en paramètres de Lamé afin de réaliser le problème direct de l'itération suivante.

Annexe B

Synthèse de la mémoire prise et du temps consommé pour les différentes classes de tests réalisées

Introduction

Nous avons synthétisé dans la Table B.1 les pics mémoires et les temps de calcul associés à chaque classe de tests. Tous les tests ont été réalisés sur le cluster de calcul JADE (Figure B.1) avec le code en Galerkin Discontinu de Brossier *et al.* (2008a), à l'exception des tests de RTM basé sur une méthode en différences finies. Pour cette dernière méthode, la règle de discrétisation à appliquer est de quatre points par longueurs d'onde ($h = \frac{\lambda}{4}$), contre 10 points par longueur d'onde pour la première. h est la taille minimum de maille à utiliser afin d'éviter les phénomènes de dispersion numérique, et λ la longueur d'onde fonction de la vitesse locale c et de la fréquence de l'onde f ($\lambda = \frac{c}{f}$). La taille de maille est donc très fine pour les tests élastiques où la plus basse vitesse de cisaillement vaut 400 m/s ; ou bien pour les tests acoustiques nécessitant une modélisation à haute fréquence ($h = 6$ m pour la RTM nécessitant de modéliser les champs jusqu'à 50 Hz). Nous n'avons pas respecté la règle de discrétisation pour le calcul des sismogrammes élastiques afin d'économiser du temps de calcul, la maille de taille 4.8 m correspondant à un échantillonnage de 8.3 points par longueur d'onde à 10 Hz. La taille des modèles est 18 km x 5.5 km pour les tests acoustiques, et de 17.3 km x 3.75 km pour les tests élastiques afin d'épargner du temps de calcul. Le milieu étant discrétisé sur un maillage hybride P0-P1, le nombre de degrés de liberté est égal à : $1.\#P0 + 3.\#P1$. Dans le cas acoustique, nous exploitons la réciprocité des fonctions de Green et traitons les sources comme des capteurs et vice-versa. Pour la ligne 21 constituée de 220 sources, cela implique 220 substitutions directes/inverses pour le calcul des champs multi-sources. Dans le cas élastique, nous n'exploitons pas la réciprocité car les capteurs disposent de deux composantes ce qui conduirait à $220 \times 2 = 440$ sources, ce qui est supérieur au nombre de tirs (320).

Calcul de la mémoire totale

Je vais maintenant expliciter la manière dont fut calculée la mémoire totale prise par chaque test. Celle-ci est la somme de trois principaux termes, nous omettons par exemple la mémoire prise par les modèles fournis au code. Le premier terme est la mémoire prise lors de la factorisation LU, associée aux communications et aux nombres de facteurs LU à stocker. Nous

SYNTHÈSE DE LA MÉMOIRE PRISE ET DU TEMPS CONSOMMÉ POUR LES DIFFÉRENTES CLASSES DE TESTS RÉALISÉES

l'obtenons en lisant la ligne dans le script de sortie de MUMPS : *EFF Min : Avg. Space in MBYTES per working proc*, à multiplier par le nombre de processus MPI utilisé pour avoir la mémoire totale associée à cette phase. Le deuxième terme associé à la phase de substitution vaut : $\#$ (nombre) de degrés de libertés \times 5 champs (vitesses + contraintes) \times 1 (champ incident puis champ rétropropagé) \times 16 (substitution par paquet de 16 sources) \times 8 octets (mémoire nécessaire au stockage d'un nombre complexe en simple précision). Le troisième terme, rempli au cours de la phase de substitution de MUMPS pour le calcul du gradient à partir des champs du second ordre, est donné par l'expression : $\#$ de degrés de libertés \times 2 champs ($V_x + V_z$) \times 2 (incidents + adjoint) $\times N_S$ (nombre de sources) \times 8 octets. Lorsque la pression est inversée, il faudrait rajouter la mémoire prise par le stockage du champ de pression aux récepteurs. La formule condensée donnant la mémoire totale (en simple précision) est la somme de : Mémoire pour la LU + $\#$ de degrés de libertés (en millions) \times $[640 + 32 \times N_S] / 1000$. Lors du calcul des sismogrammes élastiques qui est notre application nécessitant le plus de mémoire, les trois termes valent environ : 572 Gb, 4.5 Gb et 71.7 Gb respectivement. Nous avons essayé de trouver des configurations en terme de nombre de noeuds et de processus MPI qui est un compromis entre la mémoire nécessaire et le temps de calcul consommé.

Temps de calcul et multi-threading

Pour compenser le nombre limité de processus MPI lors des phases de substitutions à seconds membres multiples et du calcul du gradient, nous utilisons 2 "threads" lors des phases de substitutions effectuées avec une version "threadée" de la librairie BLAS3 (produit matrice-matrice correspondant au traitement de paquets de tirs). Si N_S , N_R et N_C représentent respectivement le nombre de sources, de récepteurs et de composantes de données, il est effectué pour la construction du gradient :

- une factorisation LU et une phase de N_S substitutions afin de résoudre le problème direct $\mathbf{A} \mathbf{p} = \mathbf{s}$, et ainsi obtenir les champs incidents \mathbf{p}_i .

- une phase de N_S substitutions, pour obtenir les champs *adjoints* rétropropagés.

Lors du calcul de la diagonale du Hessien approximé,

- est refait une phase de N_S substitutions associées au problème $\mathbf{A} \mathbf{p} = \mathbf{s}$

- est fait une phase de $N_C \cdot N_R$ substitutions associées au problème $\mathbf{A} (\mathbf{R}_d \mathbf{A}^{-1}) = \mathbf{I}$, où $(\mathbf{R}_d \mathbf{A}^{-1})$ est formé des N_R colonnes associées à chaque récepteur, afin d'obtenir les fonctions de Green aux récepteurs R_i .

Lors du calcul de la phase du pas de descente sont faits :

- une factorisation LU et une phase de N_S substitutions associées au problème $\mathbf{A}(\mathbf{m}_1) \mathbf{P} = \mathbf{S}$ afin de calculer la fonction coût obtenue dans le modèle \mathbf{m}_1 .

- une factorisation LU et une phase de N_S substitutions associées au problème $\mathbf{A}(\mathbf{m}_2) \mathbf{P} = \mathbf{S}$ afin de calculer la fonction coût obtenue dans le modèle \mathbf{m}_2 .

Il est donc fait en tout à la première itération (pour laquelle le Hessien approximé est estimé) 3 factorisations LU et $(5 N_S + N_C \cdot N_R)$ phases de substitutions. Nous allons maintenant comparer les temps nécessaires à ces opérations en prenant l'exemple d'une inversion élastique réalisée à partir des données de géophones (de la ligne 29), où trois fréquences entre 4 et 4.6 Hz sont inversées au premier groupe de fréquence, et pour lequel nous utilisons un maillage de 2.8 millions de cellules. Le temps pour la phase de substitution associée à une seule source est d'environ 1 s, contre 110 s environ pour la phase de factorisation LU. Le temps nécessaire aux 9 factorisations LU (3 LU par fréquence) pour cette première itération sera donc de 990 s. Pour ce test, nous considérons 310 sources, 195 récepteurs et nous inversons les deux composantes

de vitesses. Il sera donc fait $3 \times (5 \times 310 + 2 \times 195) = 5820$ substitutions, prenant donc 5.8 fois plus de temps que le temps de factorisation, justifiant ainsi l'optimisation des phases de substitutions en utilisant une version threadée des bibliothèques BLAS. Pour cet exemple où 2 processus MPI sont déployés par noeud et où 1, 2 et 4 threads sont dédiées aux bibliothèques BLAS, il s'est avéré que la construction du gradient prenait 772 s, 722 s et 818 s respectivement. Ce temps est la somme du temps nécessaire aux 620 substitutions (625 s) et de l'assemblage du gradient (environ 96 s). Le surplus de temps de calcul obtenu avec 4 threads doit être lié au surcoût engendré par les communications supplémentaires. A noter qu'il est possible d'utiliser un autre niveau de 'threading' lors de la phase ultérieure du calcul du gradient (et du Hessien), dont les boucles associées sont parallélisées avec OpenMP. Nous ne l'avons cependant pas utilisé en raison d'un problème d'installation sur la machine JADE. Les deux phases de calcul sont indépendantes les unes des autres si bien que ces deux niveaux de parallélisme n'interfèrent pas l'un sur l'autre. Ce niveau supplémentaire de parallélisme devrait s'avérer pour de futurs tests particulièrement utile pour l'assemblage du Hessien approximé, qui est formé par une boucle sur les cellules, les sources et les récepteurs, car nous avons remarqué que cet assemblage dans notre exemple élastique nécessite 6 fois plus de temps que les $(N_S + N_C \cdot N_R)$ substitutions effectuées pour l'obtention du Hessien.

Cas anisotrope

Dans le cas de l'approximation acoustique anisotrope, nous avons constaté que la factorisation LU devait être effectuée en double précision pour obtenir une précision suffisante tandis que la simple précision était suffisante dans le cadre de l'approximation isotrope. Cela résulte probablement du mauvais conditionnement de la matrice dans le cas acoustique-anisotrope. Rappelons que les milieux acoustiques sont par nature isotropes et que l'approximation isotrope anisotrope n'a pas de réalité physique ce qui peut expliquer ce mauvais conditionnement. Dans le cas élastique-anisotrope, nous avons conclu que la simple précision fournissait des résultats suffisamment précis. Comme l'indique la comparaison de la première et de la troisième ligne de la Table B.1, le passage en double précision coûte 2.6 fois plus en terme de mémoire et 2 fois plus en terme de temps par itération (voir colonne Durée / it.). Il est à noter que nous n'avons pas relevé ce problème en acoustique anisotrope pour les RTM réalisées avec le code en différences finies.

Cas élastique

Il est intéressant de constater dans la table B.1 qu'une inversion élastique à 6.7 Hz coûte 16 fois plus en terme de mémoire qu'une inversion isotrope, et que l'inversion de l'ensemble des fréquences demande 35 fois plus de temps CPU. Le temps de calcul nécessaire à l'inversion élastique des 5 groupes de fréquence de la composante hydrophone étant de 67,4 h (non marqué dans la Table), la réalisation du processus multi-échelle constituée de l'inversion élastique à partir de l'hydrophone puis du géophone prend en tout 6.5 jours de durée réelle, ce qui correspond à une consommation d'heures CPU de 23550 h (soit 2.7 ans...).

Remarques diverses

- Les sismogrammes acoustiques isotropes ont été calculés sur 4.096 s de temps, impliquant

SYNTHÈSE DE LA MÉMOIRE PRISE ET DU TEMPS CONSOMMÉ POUR LES DIFFÉRENTES CLASSES DE TESTS RÉALISÉES



FIGURE B.1 – Le cluster SGI Altix ICE 8200, dit JADE, est un supercalculateur scalaire parallèle d’une puissance crête de 267 Tflop/s. Ce cluster comprend 23040 cœurs répartis sur 2880 nœuds, chacun disposant de 2 processeurs Intel Quad-Core E5472 ou X5560. Chaque nœud dispose de 30 Go de mémoire utile. Le refroidissement est assuré par un système haut rendement air/eau.

un échantillonnage des fréquences de 0.24 Hz, alors que les sismogrammes dans les autres approximations ont été calculés sur 8.192 s, nécessitant de diviser par deux l’échantillonnage en fréquence.

- La version de MUMPS utilisée pour les tests nécessitant beaucoup de mémoire est la '4.9.2', car celle-ci permet un réglage plus facile qu’avec la '4.7.3' de la mémoire à allouer par processus MPI lors de la factorisation. Par contre, la version '4.7.3' est préférée pour le calcul des sismogrammes car un débordement mémoire a lieu avec la '4.9.2' pour les très basses fréquences (< 2 Hz), probablement en raison du terme en ω^2 (la pulsation) sur la diagonale de la matrice d’impédance.

- La phase d’analyse pour les tests élastique a systématiquement été réalisée avec l’algorithme AMD (*Approximate Minimum Degree*) (Amestoy *et al.*, 1996) et pas celui de METIS (Karypis & Kumar, 1999), car ce dernier aboutit à des solutions erronées pour de trop gros problèmes, bien qu’il soit plus efficace que le premier pour de plus petits problèmes.

		MEMORY					TIME									
Physique des ondes	Test	$f_{min} - f_{max}$ (Hz)	h_{moy} (m)	# degrés liberté (10^6)	LU (Gb)	Total (Gb)	# Nd	# Thrd / MPI	# MPI / Nd	# MPI	# it / 'gpf'	# 'gpf'	Durée / it	Durée / it . # MPI	Durée réelle totale	Durée CPU
Ac. Iso. SPL	FWI	3.5-6.7	19	0.64	15.8	20.8	4	1	6	24	10	5	0.25	6	12.5	400
	Sis.	0.25-21	7.5	4.41	92.8	127.7	4	1	2	8	1	86	0.04	0.32	3.7	118.4
Ac. Aniso. DBP	FWI	3.5-6.7	19	0.64	44.8	55.1	4	2	4	16	22	5	0.47	7.52	51.9	1660.8
	Sis.	1-13	11	1.91	292.5	323.6	15	2	3	45	1	99	0.07	3.4	7.5	900.0
El. Aniso. SPL	FWI	4-5.25	7.45	3.11	122	160.2	10	2	2	20	10	2	2.4	48	48	3840
		5.25-6.7	5.81	5.22	258	322.7	30	2	2	60	7	2	3.0	180	41.9	10056
	Sis.	1-10	4.84	7.0	571.8	659.0	30	2	2	60	1	75	0.14	8.4	10.2	2448
Ac. Aniso SPL (FD)	RTM	3.5-50	6	1.95	34.1	57.0	4	1	4	16	1	128	0.55	8.8	71	2272

TABLE B.1 – Table des pics mémoires (en Gb) et des temps de calcul (en heures) nécessaires à l'exécution des différentes catégories de tests

Signification des abréviations utilisées :

- Ac. : acoustique / El. : élastique
- Iso. : isotrope / Aniso. : anisotrope
- SPL : simple précision / DBP : double précision
- FWI : inversion des formes d'ondes
- Sis. : sismogrammes
- RTM : Reverse Time Migration
- it. : itération
- DF : Finite Difference
- h_{moy} : pas de maillage ou de grille moyen
- '#' : nombre

- LU : Mémoire totale nécessaire à une factorisation LU
- Total : Pic mémoire total
- Nd : noeud
- Thrd : threads associées aux bibliothèques BLAS3
- MPI : processus MPI
- 'gpf' : groupe de fréquence
- # MPI = (# MPI / Nd) . # Nd
- Durée / it = Durée réelle totale / (# 'gpf' . (# it / 'gpf'))
- Durée CPU = Durée réelle totale . 8 . # Nd

